

UNIVERSITÉ DU QUÉBEC

MÉMOIRE PRÉSENTÉ À
L'UNIVERSITÉ DU QUÉBEC À TROIS-RIVIÈRES

COMME EXIGENCE PARTIELLE
DE LA MAÎTRISE EN ÉDUCATION

PAR
ÉRIC FRENETTE

ESTIMATION DE LA DIMENSIONALITÉ DE
L'ÉCHELLE DE DÉSIRABILITÉ SOCIALE DE PAULHUS

JANVIER 1999

Université du Québec à Trois-Rivières

Service de la bibliothèque

Avertissement

L'auteur de ce mémoire ou de cette thèse a autorisé l'Université du Québec à Trois-Rivières à diffuser, à des fins non lucratives, une copie de son mémoire ou de sa thèse.

Cette diffusion n'entraîne pas une renonciation de la part de l'auteur à ses droits de propriété intellectuelle, incluant le droit d'auteur, sur ce mémoire ou cette thèse. Notamment, la reproduction ou la publication de la totalité ou d'une partie importante de ce mémoire ou de cette thèse requiert son autorisation.

Remerciements

Mes premiers remerciements s'adressent à mon directeur de recherche, monsieur Pierre Valois, professeur au dÉpartement des sciences de l'éducation de l'Université du Québec à Trois-Rivières, pour son soutien de tout instant, pour la qualité de ses conseils et surtout pour la confiance qu'il m'a manifestée tout au long de mes études. Son professionnalisme ainsi que son attitude ne sont pas étrangers au désir d'offrir ces remerciements.

Je veux également exprimer ma profonde gratitude à tous ceux qui ont permis de progresser dans ma recherche. J'aimerais souligner d'une façon particulière l'apport très apprécié de mes collègues du laboratoire de recherche sur les attitudes : Paul Villeneuve, Stéphane Thibodeau, Isabelle Boudreau, Serge Sévigny, Christiane Bordeleau et Michel Parenteau. De plus, comment passer sous silence les encouragements des membres de ma famille. Je ne pourrai jamais assez vous remercier. Enfin, j'exprime également toute ma reconnaissance aux membres du comité de sélection de la maîtrise en éducation qui ont cru en mes capacités.

Sommaire

La désirabilité sociale est un facteur pouvant contaminer les résultats des recherches en éducation. Cet artefact est susceptible d'intervenir lorsque les individus sont invités à répondre à des questionnaires portant sur leurs perceptions et lorsqu'ils anticipent la signification de la mesure. Cependant, l'estimation du nombre de dimensions d'un questionnaire de désirabilité sociale peut être problématique. La théorie moderne de réponses aux items fournit une approche prometteuse pour l'analyse de la dimensionalité. En ce sens, Ackerman, Junker, Nandakumar et Stout (1992) ont proposé l'utilisation conjointe des logiciels DIMTEST et NOHARM. L'échantillon choisi est composé de 686 personnes. Les analyses effectuées reproduisent la structure factorielle à deux dimensions tel que proposé par Paulhus (1984, 1987). De plus, chacune des deux dimensions respectent le faible principe d'unidimensionalité.

Table des matières

Introduction	8
Chapitre I — Contexte théorique	13
Désirabilité sociale	14
Théorie de réponses aux items	17
Monotonicité	18
Unidimensionalité	19
Indépendance locale	20
Techniques d'analyse de la dimensionalité	21
DIMTEST	21
NOHARM	22
Objectif de recherche.....	23
Chapitre II — Méthodologie	24
Sélection des participants	25
Instrument du mesure	25
Chapitre III — Analyse et discussion des résultats	28
Test d'unidimensionalité	29
Test de multidimensionalité	33
Conclusion.....	41
Références	44

Appendices	57
Appendice A — Modèle mathématique de la théorie de réponses aux items	57
Appendice B — Méthode de calcul du logiciel DIMTEST	61
Appendice C — Méthode de calcul du logiciel NOHARM.....	65
Appendice D — Questionnaire portant sur la désirabilité sociale élaborée par Paulhus, 1989	72

Liste des tableaux

Tableau

1	Sélection des items pour l'ensemble AT1 par DIMTEST.....	30
2	Sélection des items pour l'ensemble AT2 par DIMTEST.....	31
3	Sélection des items pour l'ensemble PT par DIMTEST	32
4	La statistique T du logiciel DIMTEST	33
5	Matrice des covariances résiduelles pour l'analyse du modèles à deux facteurs du BIDR	34
6	Coefficient de saturation de l'analyse confirmatoire à deux facteurs du BIDR	36
7	Coefficients de saturation associés à la sous-échelle autoduperie.....	37
8	Matrice des covariances résiduelles associées au facteur autoduperie	38
9	Coefficients de saturation associés à la sous-échelle hétéroduperie.....	39
10	Matrice des covariances résiduelles associées au facteur hétéroduperie	40

Liste des figures

Figure

1	Courbe caractéristique d'item.....	18
---	------------------------------------	----

INTRODUCTION

À l'aube du 21^e siècle, la mesure des attitudes en éducation est en plein essor. La rédaction d'items, leur révision, leur sélection, ainsi que la construction et la validation d'échelle d'attitude ont atteint un seuil avancé de développement.

La mesure des attitudes fait appel aux théories de la mesure et de l'échelonnement dont l'évolution est elle-même tributaire des progrès réalisés dans le domaine des statistiques et de l'éducation. Les contributions et les réussites dans ce champ d'étude ont marqué l'évolution des théories de formation et de mesure des attitudes. D'un point de vue global, l'importance du concept d'attitude dans la compréhension des comportements humains en éducation est bien établie (Gable & Wolf, 1993). En effet, force est de reconnaître que les interventions éducatives les plus efficaces passent d'abord par l'identification des déterminants psychosociaux qui influencent l'adoption d'un comportement. Malgré cette reconnaissance, de nombreux écueils se dressent devant la mesure des attitudes en éducation. Par exemple, plusieurs y opposent des critiques acerbes qui touchent, notamment, l'évaluation de ce concept.

En effet, des chercheurs affirment que les résultats colligés à l'aide de questionnaires sont quelques fois peu valides et biaisés, étant entre autres affectés par la désirabilité sociale (Berger, 1983; Garfield, 1983; LeVois, Ngyen, & Attkisson, 1981), c'est-à-dire la tendance à présenter des réponses qui reflètent un déni de pensées qui sont menaçantes au plan psychologique pour un individu, de même que la tendance à présenter une image favorable de soi à autrui. De fait, il est reconnu que cet artefact est susceptible d'apparaître lorsque les personnes sont invitées à répondre à des questionnaires portant sur leurs perceptions et lorsqu'elles anticipent la signification de la mesure (Campbell, 1975; Edwards, 1957; Scheirer, 1978; Webb, Campbell, Schartz, & Sechrest, 1970).

Depuis plus de 30 ans, de nombreuses recherches ont été effectuées dans le but de mieux cerner le concept de désirabilité sociale (DS). Malgré la multiplicité des définitions qui découlent

de ces travaux, il n'en demeure pas moins que la désirabilité sociale est et demeurera toujours un construit hypothétique, c'est-à-dire un concept impossible à observer directement. Pour pallier ces difficultés, des chercheurs tentent de mieux valider des instruments de mesure comme le *Balanced Desirable Responding* (BIDR) à l'aide de techniques métrologiques de plus en plus ingénieuses (Biggs, 1993; Evans & Forbach, 1982).

La plupart des recherches traitant de la DS ont été réalisées sur la base de la théorie classique des tests (TCT), où les réponses obtenues sont analysées en fonction du score total des individus. Plus spécifiquement, selon la TCT, le score d'un individu à un test correspond à la somme du score vrai de ce même individu et de l'erreur de mesure. Comme il est impossible de connaître le score vrai d'une personne, la solution consiste à minimiser l'erreur de mesure. En d'autres mots, plus on diminue l'erreur de mesure, plus on s'approche du score vrai de l'individu. L'idée derrière la plupart des techniques de mesure qui prennent assises sur la TCT consiste à réduire l'erreur de mesure et, partant, s'approcher du score vrai de l'individu.

L'analyse des patrons de réponses des candidats à un questionnaire tel le *Balanced Inventory Desirable Responding* (BIDR; Paulhus, 1984) démontre toutefois que les candidats présentent rarement des profils de réponses identiques. En effet, il est rare que deux personnes répondent de façon identique à tous les items d'un test. Il semble ainsi avantageux d'analyser les patrons de réponses pour mieux évaluer les répondants. La théorie de réponse aux items (TRI), développée par Rasch en 1960, permet justement d'analyser avec plus de raffinement les réponses générées par des personnes qui répondent à un questionnaire. En effet, en détectant les biais confinés aux items individuels, cette théorie rend le processus d'évaluation plus juste.

La TRI a été utilisée à maintes occasions depuis le début des années 1980 (De Ayala, 1993). Quelques-unes de ces applications incluent l'estimation de traits de personnalité (p. ex. estime de soi, affirmation de soi, ...), d'habiletés cognitives, des biais items/test, de même que

l'élaboration de tests adaptatifs. En éducation, les items et les tests recherchés sont ceux qui discriminent bien pour tous les degrés d'habiletés des étudiants. L'avantage majeur de la TRI comparativement à la TCT est qu'elle permet de déterminer les paramètres d'un item (difficulté, discrimination et pseudo-chance) indépendamment des caractéristiques des participants qui ont répondu au test. Toutefois, l'un des postulats de la TRI requiert que les tests ne mesurent qu'une seule dimension. Par exemple, il est préférable qu'un test sur les habiletés mathématiques ne soit pas biaisé par le degré d'habileté verbale des élèves complétant le test. En dépit de l'importance de ce dernier postulat, des travaux indiquent que les questionnaires utilisés en sciences de l'éducation sont rarement unidimensionnels (Hattie, 1985; Nandakumar, 1993). Ces études démontrent qu'un item est habituellement associé à une dimension dominante, mais que dans plusieurs cas il corrèle également, bien qu'à un moindre degré, avec une ou plusieurs autres dimensions secondaires du test.

La présente étude vise donc à vérifier si le BIDR (Paulhus, 1984) respecte l'un des postulats importants de la théorie de la mesure, soit celui d'unidimensionalité. Selon Humphreys (1984), un test est unidimensionnel si la majorité des items composant le test sont associés au même facteur dominant. Selon cette définition, il apparaît donc important de vérifier si la désirabilité sociale est le facteur dominant du BIDR. De plus, dans le cas où le questionnaire posséderait deux dimensions (autoduperie et hétéroduperie) comme le suggère Paulhus (1984, 1986), l'étape subséquente consistera à vérifier si chacune de ces dimensions respecte le postulat d'unidimensionalité. Ces analyses seront effectuées à l'aide des logiciels DIMTEST et NOHARM qui sont inspirés de la TRI.

Ce travail de recherche comprend quatre chapitres. Le premier chapitre permet de parcourir l'ensemble de la littérature se rattachant au problème étudié. Principalement, il traite des points de vue théoriques associés à la désirabilité sociale, des postulats de la TRI ainsi que des techniques d'analyse de la dimensionalité. L'objectif de recherche vient clore ce chapitre. Le

deuxième chapitre concerne la méthodologie de l'étude. De la présentation des participants et de l'instrument jusqu'au plan d'analyse des données, toutes les étapes de la recherche sont présentées afin de permettre une compréhension globale du travail effectué. Le troisième chapitre est consacré aux analyses statistiques effectuées pour procéder au contrôle empirique de l'objectif de recherche de même qu'à l'interprétation des résultats. La conclusion générale et la formulation de quelques recommandations sont présentées dans le quatrième et dernier chapitre et viennent mettre un terme à ce rapport de recherche.

CHAPITRE PREMIER
CONTEXTE THÉORIQUE

Le cadre théorique de cette recherche repose principalement sur l'utilisation de la théorie des réponses aux items à des fins d'évaluation de la dimensionalité de questionnaires ou de tests. En premier lieu, l'échelle de désirabilité sociale élaborée par Paulhus (1984, 1986) et traduite en français par Courmoyer et Sabourin (1991) sera décrite. En second lieu, il sera question de la théorie de réponses aux items (TRI), suivie par la présentation des postulats de base de la TRI, soit la monotonicité, l'unidimensionalité et l'indépendance locale. Troisièmement, la théorie à la base des algorithmes des logiciels DIMTEST et NOHARM sera décrite après une brève introduction aux méthodes d'estimation de la dimensionalité des tests. Enfin, la présentation de l'objectif général de la recherche vient clore ce chapitre.

Désirabilité sociale

Bien que la majorité des chercheurs souscrivent à l'idée que la désirabilité sociale constitue une source potentielle d'invalidité de la mesure des attitudes des étudiants envers un objet donné, peu d'études ont été publiées sur le sujet (Kalliopuska, 1992; Porrata, 1993; Watkins, 1996). En effet, la plupart des études portant sur les attitudes n'incluent pas de mesure de désirabilité sociale. Par exemple, dans une étude portant sur le climat de la classe et les méthodes d'apprentissage, Chan et Watkins (1994) rapportent que l'utilisation d'une échelle de désirabilité sociale en tant que variable contrôle aurait donné plus de valeur à leurs conclusions.

Lors des 30 dernières années, plusieurs questionnaires de désirabilité sociale ont été élaborés. Déjà en 1968, Wiggins recensait 12 échelles de ce type. D'autres échelles portant sur la désirabilité sociale ont été proposées depuis ce temps, dont l'«Approval-Motivation Scale» (Larsen, Martin, Ettinger, & Nelson, 1976), le «Self-and Other-Deception Questionnaire» (Sackeim & Gur, 1978), et le «Social Desirability Inventory» (Jacobson, Kellogg, Cauce, & Slavin, 1977). Selon Paulhus (1984), trois grandes approches théoriques sont à la base des questionnaires de désirabilité sociale retrouvés dans la littérature.

Selon la première approche, le concept de désirabilité sociale est composée de deux facteurs distincts, soit l'attribution et le déni (voir Millham & Jacobson, 1978). Un score élevé au facteur d'attribution signifie qu'une personne aurait tendance à s'attribuer des caractéristiques désirables au plan social. À l'opposé, l'obtention d'un score élevé sur la deuxième dimension signifie qu'une personne aurait tendance à ne pas s'attribuer des caractéristiques non valorisées dans la société. Pour mesurer ces deux facteurs, Millham (1974) divise l'échelle de Crowne-Marlowe (1960) en deux sous-échelles correspondant aux facteurs d'attribution et de déni. Toutefois, des études plus récentes démontrent que la validité factorielle de cette approche de mesure de la désirabilité sociale serait assez faible (Paulhus, 1984; Ramanaiah & Martin, 1980).

La deuxième approche de mesure de la désirabilité sociale découle des travaux de Frenkel-Brunswick (1939) et Meehl et Hataway (1946). Selon cette approche, le concept de désirabilité sociale serait composé des dimensions d'autoduperie et d'hétéroduperie (Paulhus, 1984, 1986; Cournoyer & Sabourin, 1991). L'autoduperie (traduction française de l'expression «Self-Deception»; c.-à-d. méconnaissance de soi ou auto-illusion) reflète un déni de pensées qui sont menaçantes au plan psychologique pour un individu et, a pour fonction de le protéger de connaissances douloureuses ou anxiogènes qu'il a sur lui-même. Quant à l'hétéroduperie (traduction française de l'expression «Impression Management»; c.-à-d. contrôle social de l'image de soi ou aveuglement d'autrui), elle désigne la tendance à présenter une image favorable de soi à autrui. La validité convergente et divergente de ces échelles ont été supportées dans une série d'études expérimentales et corrélationnelles (Paulhus, 1986).

Enfin, la troisième approche théorique suggère que le concept de désirabilité sociale posséderait une structure bi-dimensionnelle formée des facteurs «Alpha» et «Gamma» (Block, 1965; Wiggins, 1964). L'obtention d'un score élevé sur le facteur «Alpha» signifie que de façon inconsciente la personne a tendance à fournir des réponses biaisées. Plusieurs échelles sont utilisées pour mesurer cette tendance de réponse : l'échelle de désirabilité sociale de Edwards

(1957), l'échelle de répression-sensibilisation de Byrne (1961), l'échelle de facilitation-inhibition de Uhlmann (1962) et l'échelle d'attitude à l'égard de l'administration d'un questionnaire de Meehl et Hathaway (1946). À l'opposé, un score élevé sur le facteur «Gamma» indique que les personnes biaisent délibérément leurs réponses à un questionnaire (Edwards, Diers, & Walker, 1962; Jackson & Messick, 1962). Plusieurs échelles ont été proposées pour mesurer ce facteur dont l'échelle de désirabilité sociale de Wiggins (1959), la «Positive Malinger Scale» de Cofer, Chance et Judson (1949) et «Minnesota Multiphasic Personality Inventory Lie Scale» de Meehl et Hataway (1946).

Jusqu'à ce jour, il n'existerait qu'une seule étude (Paulhus, 1987) examinant les relations entre tous ces questionnaires de désirabilité sociale. L'auteur a observé que les items des questionnaires mesurant le facteur «Alpha» présentaient des coefficients de saturation élevés sur le facteur d'autoduperie. Similairement, les items des questionnaires mesurant le facteur «Gamma» présentaient des coefficients de saturation élevés sur le facteur d'hétéroduperie. Il ressort de cette étude comparative que le concept de désirabilité sociale serait composé de deux dimensions orthogonales «Autoduperie» et «Hétéroduperie». Cependant, la valeur de cette structure factorielle n'a pas été étudiée de façon exhaustive à l'aide de la théorie de réponses aux items. Cette théorie permettra d'évaluer avec plus de précision la dimensionalité de l'échelle. Cependant, il est à noter que le développement de modèles basés sur la théorie de réponses aux items et permettant l'analyse des choix de réponses polytomiques (p. ex échelle de type Likert en 5 points) n'en est encore qu'au stade embryonnaire. Dans le cadre de la présente étude, il semble donc plus approprié d'utiliser les modèles pour choix de réponses dichotomiques. En ce sens, les items du questionnaire de désirabilité sociale (BIDR) seront utilisés comme des items dichotomiques. À cet égard, Paulhus (1984) propose une méthode pour dichotomiser les items du BIDR. Les détails de cette méthode seront présentés dans la section méthodologique. Selon Schaeffer (1988) la dichotomisation des scores n'invalide pas le processus d'analyse d'items et de tests.

Jusqu'à présent, quelques chercheurs se sont intéressés à l'élaboration de méthodes visant à déterminer le nombre de dimensions présentes dans un questionnaire ou un test composé d'items dichotomiques. Parmi les méthodes les plus prometteuses, il faut sans aucun doute prendre en compte celle élaborée par Ackerman, Junker, Stout et Nandakumar en 1992. Selon ces auteurs, l'examen de la dimensionalité d'un test s'effectue en deux étapes : (a) vérification de l'hypothèse d'unidimensionalité et (b) dans le cas où l'hypothèse d'unidimensionalité serait infirmée, examen de la structure multidimensionnelle du test. Pour ce faire, Ackerman et al. (1992) ont proposé l'utilisation du logiciel DIMTEST lors de la première étape et du logiciel NOHARM lors de la deuxième étape. Puisqu'elle est à la base du développement des algorithmes des logiciels DIMTEST et NOHARM, la théorie de réponse aux items sera présentée dans les paragraphes qui suivent.

Théorie de réponse aux items (TRI)

La théorie de réponses aux items (TRI) est de plus en plus utilisée pour la construction de banques d'items de même que la conception de tests adaptatifs et la validation de questionnaires. Cette théorie moderne des tests s'avère plus efficace que la théorie classique des tests, et ce pour deux raisons principales. Tout d'abord, la TRI permet une meilleure estimation de l'habileté ou du trait mesuré chez les individus puisque l'analyse porte sur les patrons de réponses aux items plutôt que sur le score total tel que calculé traditionnellement. Deuxièmement, la TRI permet une meilleure estimation des caractéristiques métrologiques des items (c.-à-d. leur degré de difficulté, leur degré de discrimination de même que le facteur de pseudo-chance). Cette théorie prend appui sur un modèle mathématique (voir l'appendice A) que l'on peut représenter par une fonction (souvent monotone croissante) reliant l'habileté du répondant à la probabilité de réussir un item donné. Cette fonction est appelée courbe caractéristique d'item.

Les trois hypothèses de base de la TRI sont la monotonie, l'unidimensionalité et l'indépendance locale. La monotonie suppose que la probabilité de répondre correctement à un item croît en fonction de l'habileté ou du trait mesuré. Quant à l'unidimensionalité, elle suppose que les items d'un test mesurent une seule habileté. Enfin, l'indépendance locale suppose que pour un degré d'habileté particulier, les réponses données par un même répondant aux différents items sont statistiquement indépendantes entre elles. Ces trois postulats de la TRI sont explicités plus bas.

Monotonie

Une courbe caractéristique d'items (CCI) est dite monotone lorsque celle-ci prend la forme de l'ogive normale (en forme de S). En d'autres mots, une CCI est monotone lorsque les individus présentant des scores élevés par rapport à un trait donné ont une plus grande probabilité de répondre correctement à l'item (dans le cas d'un item à choix multiples) ou d'endosser une option donnée (p. ex. dans le cas d'un item de type Likert) que les personnes présentant des scores faibles. La Figure 1 illustre bien ce concept de monotonie.

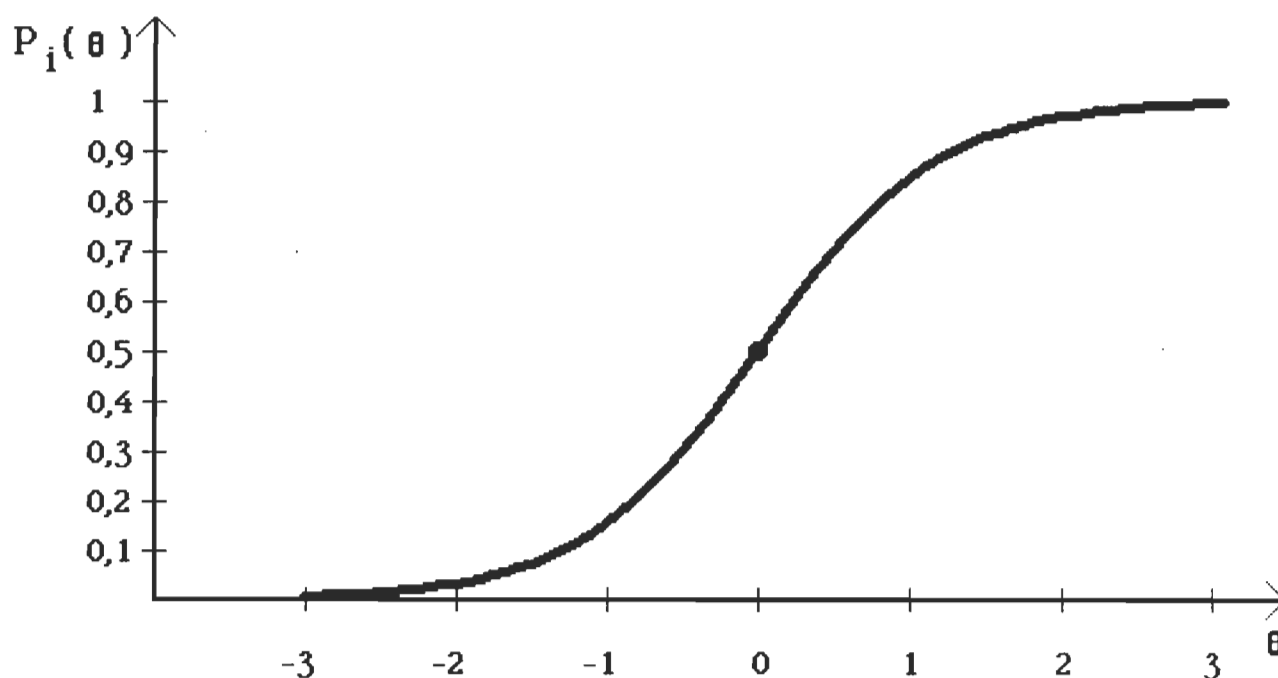


Figure 1. Courbe caractéristique d'item

Unidimensionalité

Le concept d'unidimensionalité est au coeur même des recherches menées en docimologie (Nandakumar, 1993; Zwick, 1987). Selon Hattie (1984, 1985), un test est unidimensionnel si un seul facteur est nécessaire pour rendre compte du patron de réponses des répondants à un test particulier. Cette définition est basée sur la théorie des traits latents. Pourtant même si dès le début des années 30 l'unidimensionalité était considérée comme une caractéristique que tout test devait posséder (Walker, 1931), plusieurs chercheurs soutiennent que ce postulat de la théorie de la mesure n'est que rarement éprouvée (Hambleton, Swaminathan, Cook, Eignor, & Gifford, 1978; Hattie, 1984, 1985; Lord 1980; Lord & Novick 1968; Nandakumar & Stout, 1993; Stout, 1987, 1990). À cet égard, Lord (1980) affirme : «Il existe un grand besoin de créer un test statistique pour éprouver l'hypothèse d'unidimensionalité d'un test» (traduction libre, p. 21). Pour leur part, Hambleton et al. (1978) mentionnent que l'hypothèse d'unidimensionalité doit être vérifiée avant d'estimer les coefficients d'adéquation des modèles de trait latent (p. ex. modèles à 1, 2 ou 3 paramètres).

Au cours des années 80 et 90, plusieurs auteurs proposent des modèles visant à estimer le nombre de dimensions que possède un test. Dans une recension de ces modèles, Hattie (1984, 1985) évalue théoriquement et empiriquement plus de 80 indices d'unidimensionnalité. Suite à l'examen de ces indices, il conclut en la médiocrité de la plupart de ceux-ci. Ses études démontrent que les indices définis en terme de rang unique (Lumsden, 1957), du pourcentage de variance expliquée par la première composante ou facteur (Lord, 1980), ou en terme de dérogation à l'échelle parfaite de Guttman (Guttman, 1944, 1950), du type de matrice de corrélations utilisée (Hulin, Drasgow, & Parsons, 1983), ou du nombre de facteurs communs (Tucker & Lewis, 1973) ne constituent pas des estimés valides de l'unidimensionalité. Il appert que seuls les indices développés sur la base du postulat d'indépendance locale (Lord & Novick,

1968; McDonald, 1981) et du modèle factoriel non linéaire (Goldstein, 1980; Hattie, Krakowski, Rogers, & Swaminathan, 1996; McDonald & Ahlawat, 1974) s'avèrent efficaces pour tester l'unidimensionnalité des tests. En effet, selon ces derniers auteurs l'utilisation de l'analyse factorielle non linéaire est préférable à l'analyse factorielle linéaire pour le traitement de données dichotomiques puisque l'analyse factorielle linéaire produit des matrices non grammiennes (non positive définit).

Par sa part, Stout (1987) mentionne que l'utilisation d'un test statistique pour confirmer ou infirmer l'hypothèse d'unidimensionalité peut être utile pour : (a) éviter qu'un test mesurant un trait donné ne soit contaminé par un autre trait, (b) diviser un test en deux ou plusieurs parties lorsque celui-ci mesure deux ou plusieurs traits latents, et ce même si les items qui composent le test ont un certain lien avec ce concept, et (c) s'assurer de respecter le postulat d'unidimensionalité avant d'utiliser les modèles unidimensionnels de la TRI.

Indépendance locale

Il est largement reconnu que la majorité des examens, tests ou questionnaires possède plusieurs dimensions (Hambelton & Swaminathan, 1985; Humphreys, 1981, 1985, 1986; Reckase, 1979, 1985; Stout, 1987; Traub, 1983; Yen, 1985). Ces dimensions peuvent être considérées comme dominantes ou secondaires. Une dimension est dite dominante lorsque presque tous les items d'un test y sont associés assez fortement (Stout, 1990). À l'opposé, les dimensions secondaires ne sont constituées que d'un sous-ensemble d'items plus fortement liés entre eux qu'avec les autres items du test. Traub (1983) mentionne que les antécédents scolaires de même que les stratégies d'apprentissage utilisées représentent des exemples de dimensions secondaires pouvant affectées un test.

L'indépendance locale suppose que pour un niveau d'habileté ou de trait latent donné, les réponses d'un individu aux items sont statistiquement indépendantes entre elles, et ce autant pour

les dimensions dominantes que secondaires d'un test. Ce postulat d'indépendance locale est très difficile à respecter. En effet, pour chaque dimension mesurée par le test, il requiert que les covariances entre items soient égales à 0 et que les réponses obtenues aux items suivent la loi de densité normale multivariée. Étant donné la difficulté de rencontrer ce postulat d'indépendance locale, McDonald (1979) suggère de le remplacer par le postulat faible d'indépendance locale. Ce postulat requiert également que les covariances entre les items soient égales à 0, mais n'implique pas que les réponses obtenues aux items suivent la loi de densité normale multivariée.

Techniques d'analyse de la dimensionalité

Pour tester la dimensionalité du BIDR, deux algorithmes de calcul seront employés. Le programme DIMTEST permet de vérifier le postulat d'unidimensionalité. Quant au programme NOHARM, il est utilisé pour vérifier le nombre de dimensions sous-jacent à un test dans le cas où l'hypothèse d'unidimensionalité est infirmée.

DIMTEST

Bien qu'en général le principe d'indépendance locale soit théoriquement et mathématiquement accepté, Stout (1987, 1990) et Humphreys (1984) prétendent qu'il est en pratique difficile à respecter. En effet, selon ces auteurs chacun des items d'un test relève non pas d'une seule dimension mais de plusieurs dimensions. Pour pallier ce problème, Stout (1987, 1990) propose le postulat d'indépendance essentielle qui constitue une variante moins restrictive du postulat d'indépendance locale.

Contrairement au postulat d'indépendance locale qui requiert que les covariances entre items soient égales à 0 pour chaque dimension (dominantes et secondaires) mesurée par le test, le postulat d'indépendance essentielle nécessite que ces covariances inter-items tendent vers 0, et ce seulement pour les dimensions dominantes. Ce postulat d'indépendance essentielle est à la base même de l'algorithme du logiciel non paramétrique DIMTEST (pour plus de plus amples détails

en regard des développements théoriques voir : Junker, 1990, 1991, 1992; Nandakumar, 1987, 1991; Nandakumar & Stout, 1993; Stout, 1990; Stout et al., 1991). Lors de la dernière décennie, des études ont démontré la robustesse de la statistique T du DIMTEST élaboré par Stout (Hattie et al., 1996; Nandakumar, 1993; Stout, 1990). Les quatre étapes du logiciel DIMTEST sont présentées à l'annexe B.

Dans une étude de type Monte carlo comparant quatre techniques d'analyse de la dimensionalité des tests (DIMTEST, NOFA, la procédure de Holland, l'analyse factorielle linéaire), Nandakumar (1994) démontre que le DIMTEST s'avère le plus efficace pour éprouver l'unidimensionalité d'une banque de données. Par contre, ce dernier obtient de meilleures performances avec un test de plus de 25 items et la présence d'au moins de 500 sujets (De Champlain & Gessaroli, 1991; Stout, 1990).

Rappelons que le logiciel DIMTEST permet de vérifier si un test est unidimensionnel ou non. Toutefois, dans le cas où le postulat d'unidimensionalité est rejeté il ne donne aucune information quant au nombre de dimensions composant le test. Hattie (1984) ainsi que De Champlain et Gessaroli (1991) soulignent l'efficacité de la technique d'analyse factorielle non linéaire, telle qu'utilisée par le logiciel NOHARM, pour évaluer le nombre de dimensions d'un test. Ackerman et al. (1992) suggèrent d'utiliser de façon complémentaire les logiciels DIMTEST et NOHARM pour examiner la dimensionalité d'un construit.

NOHARM

Le modèle normal ogive (c.-à-d. la courbe des fréquences cumulées) illustré à la Figure 1 représente une courbe en forme de S. Plusieurs auteurs démontrent que ce genre de courbes ne peut être représenté par un modèle linéaire, suggérant ainsi que le modèle d'analyse factorielle linéaire pourrait ne pas évaluer avec précision le nombre de dimensions d'un test. McDonald (1967, 1982) démontre que l'analyse factorielle non linéaire permettait d'évaluer avec justesse

les courbes en forme de S, et partant, le nombre de dimensions qui composent un test construit selon les postulats de la théorie de réponses aux items (Goldstein, 1980; Goldstein & Wood, 1989; Hattie 1984, 1985; Traub & Lam, 1985). McDonald (1967a) ainsi que Takane et De Leeuw (1987) démontrent même que les modèles de la TRI sont des cas particuliers de l'analyse factorielle non linéaire.

S'inspirant de l'analyse factorielle non linéaire, McDonald (1981, 1983, 1985) propose une nouvelle approche pour l'analyse de la dimensionalité des tests. Suite aux travaux de McDonald, Fraser (1988) opérationnalise cette approche sous la forme du logiciel NOHARM (voir l'annexe C). L'un des atouts du programme NOHARM est qu'il permet l'analyse exploratoire et confirmatoire des modèles multidimensionnels. De plus, il fournit la matrice des covariances résiduelles d'items comme indice d'adéquation du modèle unidimensionnel ou multidimensionnel testé. Plusieurs études ont utilisé avec succès le logiciel NOHARM pour éprouver la dimensionalité d'un test (Ackerman, 1996; Hambleton & Rovinelli, 1986; Hattie, 1984; Knol & Berger, 1991; Nandakumar, 1991).

Objectif de recherche

Dans le cadre de la présente étude, il est postulé que le *Balanced Inventory Desirable Responding* est composé des dimensions autoduperie et hétéroduperie tel que proposé par Paulhus (1986). Il est également postulé que les deux dimensions sont unidimensionnelles. Pour vérifier ces hypothèses de recherches, nous aurons recours à des approches métrologiques basée sur la théorie moderne de réponses aux items.

CHAPITRE DEUXIÈME

MÉTHODOLOGIE

Ce chapitre traite de l'ensemble des opérations déployées pour réaliser la présente recherche. Les différentes parties de la méthode y sont présentées : (a) la sélection des participants, (b) l'instrument de mesure utilisé et enfin, (c) le plan d'analyse des résultats.

Sélection des participants

L'échantillon initial se compose de 1372 adultes. L'âge des participants varie entre 18 et 56 ans; la moyenne d'âge des femmes et des hommes est respectivement de 29.94 (E.T. = 4.67) et 32.05 (E.T. = 5.55) ans. Les salaires des hommes varient entre 0\$ et 700,000\$. Leur revenu moyen est de 34,782\$ (médiane = 30,000\$) avec un écart type de 31,834\$ (indice interquartile = 16,000\$). Pour les femmes, les salaires oscillent entre 0\$ et 140,000\$. Leur revenu moyen est de 20,108\$ (médiane = 20,000\$) avec un écart type de 15,768\$ (indice interquartile = 17 000\$). Les participants, hommes et femmes, ont en moyenne 13 ans de scolarité.

Instrument de mesure

L'inventaire de désirabilité sociale (*Balanced Inventory of Desirable Responding*), mis au point par Paulhus (1984) et validé en version française par Sabourin, Laferrière, Sicuro, Coallier, Cournoyer et Gendreau (1989), comprend 40 items et est coté à partir d'une échelle en sept points de type Likert allant de faux (+1) à totalement vrai (+7). Ce questionnaire définit la désirabilité sociale en fonction de deux facteurs : l'autoduperie et l'hétéroduperie. Il est inspiré des deux questionnaires élaborés par Sackeim et Gur (1978); le *Self-Deception Questionnaire* et le *Other Deception Questionnaire*.

La sous-échelle d'autoduperie, composée de 20 items, évalue le déni de pensées qui sont psychologiquement menaçantes tout en étant supposément universellement vraies : la haine d'un de ses parents, le plaisir ressenti à faire ses selles, la peur de l'homosexualité. Seuls les répondants peuvent confirmer la véracité ou la fausseté de leurs réponses. Les scores vont de 20 à

140 (0 à 20 selon l'échelle de scores dichotomiques), les scores les plus élevés reflétant les plus hauts niveaux d'autoduperie.

D'autre part, la sous-échelle d'hétéroduperie est composée de 20 items et contient des énoncés qui sont socialement désirables mais qui désignent des comportements peu fréquents : «Je dis toujours la vérité», «J'ai déjà reçu trop de monnaie d'un(e) caissier(ère) sans le lui dire», «Je respecte toujours le code de la route même s'il est peu probable que je me fasse arrêter». Les scores oscillent entre 20 et 140 (entre 0 et 20 pour l'échelle de scores dichotomiques), les scores les plus élevés reflétant un plus haut niveau d'hétéroduperie.

Les validités convergentes et divergentes de ces sous-échelles sont supportées par une série d'études expérimentales et corrélationnelles (Gur & Sackeim, 1979; Paulhus, 1982, 1984; Paulhus & Levitt, 1987; Sackeim & Gur, 1978, 1979). Le BIDR a été traduit en langue française par Sabourin, Lussier, Courmoyer, Lecours et Wright (1988). Les coefficients de cohérence interne sont de .75 pour l'autoduperie et de .70 pour l'échelle d'hétéroduperie. La validité de cette traduction a été éprouvée en corrélant les cotes d'autoduperie et d'hétéroduperie à des scores de détresse psychologique (Sabourin, Laferrière, Sicuro, Coallier, Courmoyer, & Gendreau, 1989). Les résultats obtenus confirment ceux d'autres études (Linden, Paulhus, & Dobson, 1986). Dans le cadre de la présente étude les coefficients alpha atteignent .72 pour l'échelle d'autoduperie et de .67 pour l'échelle d'hétéroduperie.

Le but de cette recherche consiste à vérifier l'unidimensionalité d'une échelle de mesure portant sur la désirabilité sociale. En cas de rejet de cette hypothèse, il devient alors important de vérifier le nombre de groupements unidimensionnels. La première étape consiste donc à éprouver le concept d'unidimensionalité pour l'ensemble de l'échelle de désirabilité sociale. Cette étape sera effectuée sur la base du postulat de la dimensionalité essentielle (DIMTEST). En cas de rejet de l'hypothèse d'unidimensionalité essentielle, il est important de vérifier le nombre de

dimensions composant le BIDR. Le logiciel NOHARM permettra de procéder à cette analyse. Par la suite, l'undimensionalité sera vérifiée pour chacune des deux dimensions postulées par Paulhus (1984, 1986, 1991). Cette étape aura recours au logiciel NOHARM plutôt qu'au logiciel DIMTEST qui nécessite la présence d'au moins 25 items. Les logiciels DIMTEST et NOHARM sont décrits plus en spécifiquement dans les appendices B et C.

Il est cependant à noter que les logiciels utilisés (NOHARM et DIMTEST) nécessitent l'utilisation de données dichotomiques. Les résultats obtenus sont ainsi transformés en données dichotomiques selon les recommandations de Paulhus (1991) : les valeurs 6 et 7 prennent maintenant la valeur 1 pour indiquer la présence de désirabilité sociale, tandis que les valeurs 1 à 5 prennent la valeur 0 pour signaler l'absence de désirabilité sociale (voir l'appendice D pour plus d'informations).

CHAPITRE TROISIÈME
ANALYSE ET INTERPRÉTATION DES RÉSULTATS

Ce chapitre qui traite de la dimensionalité du BIDR se divise en deux parties. La première partie vise à vérifier si cet inventaire de désirabilité sociale possède ou non un facteur dominant (unidimensionalité). Deuxièmement, si le BIDR est composé des dimensions autoduperie et hétéroduperie comme le suggère Paulhus (1984, 1986), il s'agit alors de vérifier si chacune de ces dimensions respecte à leur tour le postulat d'unidimensionalité. Enfin, étant donné la nature méthodologique de cette recherche, il est apparu préférable de procéder simultanément à la description et à l'interprétation des résultats.

Test d'unidimensionalité

L'échelle de désirabilité sociale est testée en premier lieu pour éprouver l'hypothèse d'unidimensionalité à l'aide du logiciel DIMTEST. Selon Stout (1990), le postulat d'unidimensionalité se vérifie à travers de quatre étapes. La première étape du DIMTEST consiste à choisir un sous-ensemble d'items qui démontrent le plus d'homogénéité entre eux et le moins de ressemblance conceptuelle avec les autres items du questionnaire (regroupement d'items nommé AT1 par le concepteur du logiciel). Dans le cas où ce regroupement d'items particuliers ne se différencie pas des autres items, il est alors permis de croire que le postulat d'unidimensionalité est respecté. Dans le cas contraire, il faut supposer que le questionnaire est composé de plusieurs dimensions. Stout et al.(1996) proposent trois stratégies pour sélectionner les items de l'ensemble AT1 : (1) l'analyse factorielle exploratoire; (2) l'analyse de contenu par des experts; (3) l'analyse par grappes («cluster analysis»).

Dans la présente recherche, l'analyse factorielle exploratoire est la stratégie qui est privilégiée pour choisir les items de l'ensemble AT1. Cette analyse est effectuée auprès de 500 des 1366 participants de l'étude. Les réponses fournies par les autres participants ($n = 866$) sont

utilisées pour leur part aux étapes 2, 3 et 4. Suite à cette analyse et tel que recommandé par Nandakumar (1993), les cinq items démontrant les plus hauts coefficients de saturation sur le second facteur forment l'ensemble AT1 (voir le Tableau 1). Le test de Wilcoxon (Wilcoxon rank sum test) est ensuite utilisé pour vérifier si le degré de difficulté des items de l'ensemble AT1 est trop bas. Les résultats indiquent que les items ne présentent pas des indices de difficulté trop bas, $p = .90$.

La deuxième étape consiste à créer un deuxième ensemble, nommé AT2, (voir le Tableau 2) dont les items présentent des similarités avec les items de l'ensemble AT1, c'est-à-dire que leur niveau de difficulté soit du même ordre de grandeur et qu'ils présentent peu de ressemblance conceptuelle avec le reste des items du test. Cet ensemble d'items permettra de corriger la statistique T d'unidimensionalité calculée à l'étape 4. Le choix de l'ensemble AT2 est effectué par l'algorithme de DIMTEST.

Tableau 1
Sélection par DIMTEST des items de l'ensemble AT1

Item	Coefficient de saturation du 2 ^e facteur	Niveau de difficulté de l'item	Rang selon le niveau de difficulté
30	.61	.26	7
31	.47	.17	3
26	.39	.46	20
3	.34	.27	8
29	.32	.68	33

La troisième étape ($n = 866$) consiste à grouper le reste des items dans un nouvel ensemble d'items nommé PT (voir le Tableau 3). Par la suite, les scores obtenus à ce groupe

d'items par les participants sont additionnés pour créer un score total. Finalement, tel que recommandé par Stout (1987), les répondants présentant des scores totaux similaires sont regroupés de façon à former 20 groupes. Dans la présente étude, 15 des 20 groupes formés possèdent un nombre suffisant de personnes (20) pour effectuer le calcul de la statistique \underline{T} conservatrice de Stout (1987, 1990) de même que celui de la statistique \underline{T} plus puissante de Nandakumar (1992) (quatrième étape). Donc, l'analyse d'unidimensionalité du BIDR avec le logiciel DIMTEST est effectuée sur 89.72% des 866 participants ($n = 777$). Ce pourcentage est supérieur au seuil (85%) suggéré par Nandakumar (1992).

Tableau 2
Sélection par DIMTEST des items de l'ensemble AT2

Item	Coefficient de saturation du 2 ^e facteur	Niveau de difficulté de l'item	Rang selon le niveau de difficulté
22	.22	.10	2
16	-.16	.38	9
10	.15	.38	100
38	-.18	.48	21
1	.02	.66	32

Le calcul de la statistique \underline{T} conservatrice de Stout (1987, 1990) et de la statistique \underline{T} plus puissante de Nandakumar (1992) constitue la quatrième étape du logiciel DIMTEST. Elle consiste à vérifier l'unidimensionalité du test ou du questionnaire (voir l'appendice B pour plus de détails au sujet du calcul de ces valeurs \underline{T}). Dans les deux cas, DIMTEST calcule d'abord une statistique \underline{T} pour chacun des ensembles AT1 et AT2 à partir des groupes de personnes formés à l'étape 3, soit \underline{T}_L et \underline{T}_B respectivement. Par la suite ces deux valeurs sont utilisées pour calculer la statistique non biaisée \underline{T} .

Tableau 3
Sélection par DIMTEST des items de l'ensemble PT

Item	Coefficient de saturation du 2 ^e facteur	Niveau de difficulté de l'item	Rang selon le niveau de difficulté
2	-.24	.62	27
4	-.37	.73	34
5	-.11	.41	12
6	.29	.65	31
7	-.12	.36	11
8	.06	.65	30
9	-.19	.42	16
11	-.04	.21	5
12	.00	.41	14
13	-.04	.07	1
14	.06	.46	19
15	-.13	.59	25
17	.05	.46	18
18	-.30	.43	17
19	.17	.64	29
20	.02	.42	15
21	-.09	.50	22
23	-.39	.51	23
24	-.40	.74	36
25	.25	.22	6
27	-.14	.78	37
28	-.15	.63	28
32	.22	.60	26
33	.04	.41	13
34	.08	.73	35
35	-.43	.18	4
36	.22	.87	38
37	.24	.52	24
39	-.42	.89	39
40	-.15	.90	40

Les résultats obtenues démontrent que le BIDR ne possède pas une dimension dominante : \underline{T} conservatrice = 5.83, $p < .001$ et \underline{T} plus puissante = 6.95, $p < .001$ (voir le Tableau 4). Cependant, Sout (1990) recommande de recommencer plusieurs fois la même série d'analyses en sélectionnant à chaque occasion différents sous-groupes de participants pour choisir les items

de l'ensemble ATI. La séquence d'analyse a donc été reprise une quarantaine de fois. À chaque fois les résultats indiquent que le BIDR n'est pas unidimensionnel.

Deux recommandations peuvent être formulées à partir de ces résultats. Premièrement, d'un point de vue théorique, les présents résultats suggèrent que les analyses d'items effectuées sur la base de modèles unidimensionnels de la TRI (p. ex. modèle ogive et modèles à 1, 2 ou 3 paramètres) doivent être effectuées sur les items de chacune des dimensions du BIDR plutôt que sur l'ensemble des items de l'inventaire. Deuxièmement, d'un point de vue pratique, les résultats indiquent que la stratégie qui consiste à effectuer des analyses statistiques sur le score total du BIDR peut mener à des erreurs de classement des répondants. Par conséquent, il est suggéré aux utilisateurs du BIDR d'effectuer leurs analyses sur les sous-dimensions du BIDR plutôt que sur l'ensemble du questionnaire.

Tableau 4

Les valeurs \underline{T} conservatrice et plus puissante obtenues pour évaluer l'unidimensionalité du BIDR

Méthode de calcul du \underline{T}	\underline{T}_c	\underline{T}_p	\underline{T}	p
\underline{T} conservatrice de Stout	5.83	.17	4.00	.00003
\underline{T} plus puissante de Nandakumar	6.95	.10	4.84	.00001

Test de multidimensionalité

Les résultats précédents démontrent que le BIDR ne représente pas un concept unidimensionnel. Cette deuxième partie du chapitre consiste donc à évaluer le nombre de dimensions composant ce questionnaire de désirabilité sociale. L'analyse factorielle non linéaire et confirmatoire effectuée à partir du logiciel NOHARM permettra d'éprouver le modèle à deux dimensions (autoduperie et hétéroduperie) postulé par Paulhus (1986).

Deux stratégies sont retenues pour vérifier le degré d'ajustement du modèle. La première stratégie consiste à examiner la matrice des covariances résiduelles des items du modèle (c.-à-d. les carrés moyens résiduels) afin de vérifier le postulat d'indépendance locale. Plus spécifiquement, selon Fraser (1988) l'examen de ces variances résiduelles permet d'identifier les items qui ne s'ajustent pas au modèle et de ce fait la présence d'une ou de plusieurs autres dimensions que celles proposées dans le modèle testé. À cet égard, l'examen approfondi de la matrice des covariances résiduelles (Tableau 5) indique la présence d'aucune valeurs résiduelles élevées, ce qui suggère que le BIDR serait tel que prévu formé des dimensions d'autoduperie et d'hétéroduperie.

Tableau 5

Matrice des covariances résiduelles pour l'analyse du modèle à deux facteurs

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
2	.004																			
3	-.002	.004																		
4	.002	-.027	.017																	
5	.012	-.006	.011	-.023																
6	-.029	-.006	-.010	.011	.007															
7	-.004	-.005	.003	.000	-.011	-.008														
8	.019	-.007	-.024	.009	-.013	.003	-.008													
9	.006	-.003	.017	-.023	-.014	.016	-.011	.005												
10	-.015	.002	-.013	.022	.014	.013	.007	-.010	.015											
11	.002	.007	.006	.000	-.005	.004	-.017	-.006	-.012	0.12										
12	-.004	-.013	-.003	.007	.013	.006	.000	.012	.007	-.008	.005									
13	.004	.000	.002	.001	.000	.006	-.005	-.002	-.007	.003	-.007	.006								
14	-.006	.014	.012	.004	-.040	-.030	.004	.000	.007	-.009	.000	-.002	.003							
15	-.006	.004	-.002	-.002	.006	.002	.009	.012	-.013	.011	-.020	.010	-.006	.006						
16	.002	.007	.004	.002	.012	.011	.004	.011	.004	-.002	.011	-.027	.005	.003	-.010					
17	-.010	.010	-.004	.008	.009	-.009	-.010	.013	.000	.008	.011	-.012	-.001	.004	-.009	-.015				
18	.022	-.008	.010	-.013	-.007	.028	.007	-.017	-.005	.011	-.010	-.002	-.006	.015	-.012	-.004	.008			
19	-.009	.011	-.005	.009	.020	-.010	.012	.012	.009	-.008	.008	.000	.003	.005	.009	-.010	-.012	.016		
20	-.013	-.007	.000	-.013	.016	-.002	.000	.013	-.002	.009	.016	-.023	.008	.007	.008	-.014	-.009	.005	-.015	
21	.013	-.011	.009	-.004	.003	.021	-.007	-.003	-.004	.005	.004	.005	.000	.007	-.001	.013	-.001	-.002	.003	.011
22	.000	.001	.000	.000	-.002	-.002	-.002	.000	.006	.000	-.005	-.005	.000	-.002	.004	-.001	-.002	-.001	.001	.001
23	.003	-.004	.011	-.018	-.016	.018	.004	.000	-.018	.006	-.010	-.003	.000	-.006	.009	-.014	.011	-.022	.002	-.005
24	.000	-.002	.008	-.017	-.005	.008	-.004	.015	-.012	.012	-.006	.005	-.002	.009	.000	-.008	.005	-.011	-.006	-.004
25	.003	.005	-.010	.009	.004	-.007	.000	-.001	.002	-.008	-.010	.000	-.010	.004	-.007	.000	-.001	-.005	.002	.004
26	.000	.008	-.013	.019	.005	-.014	.006	-.012	.006	-.006	-.008	.000	-.003	-.006	-.004	.001	-.007	.019	-.027	.002
27	.003	.000	.003	-.004	.006	-.003	.005	-.009	-.005	.004	.002	.003	-.001	.001	-.002	.008	.011	-.013	.004	.000
28	.003	-.002	.000	-.009	-.002	.010	.004	-.009	-.005	-.005	-.005	.005	.002	.010	-.010	.003	-.021	.004	-.003	
29	-.010	-.005	-.014	.007	.014	-.032	.003	.000	.003	-.006	.002	.006	.002	-.004	.004	.000	.000	.018	-.022	.003
30	.005	.008	-.017	.023	.009	-.001	.008	-.014	.012	-.006	-.003	-.003	-.002	.000	.000	.008	.003	.003	-.011	.000
31	.004	-.002	-.023	.009	.008	.005	.011	-.013	.000	-.003	-.001	-.002	-.002	.016	-.005	.003	.006	-.001	.000	.002
32	-.003	.011	-.004	.012	.014	.002	.005	.013	.012	-.007	.011	-.001	.003	.006	-.007	.004	-.024	.009	-.010	.003
33	-.008	-.004	-.006	.000	-.004	-.004	-.002	-.019	.003	-.017	-.010	.010	-.001	-.005	.003	.019	.009	-.001	.009	-.003
34	-.001	.011	.002	.011	.013	.002	.012	.016	.003	-.011	.000	.013	.004	.002	-.003	.004	-.003	.005	-.013	.010
35	-.005	.006	.007	-.001	.003	.015	-.007	-.005	-.017	.000	-.002	-.007	-.008	.008	-.005	-.008	-.004	-.019	.012	.002
36	.001	-.002	-.009	.000	.003	.000	.007	-.002	.011	.000	.006	.005	.004	.005	-.001	.000	-.001	-.004	.004	-.003
37	.010	.003	-.005	.007	-.001	.003	.007	.005	.002	-.009	-.001	.003	-.004	-.010	-.009	-.006	.008	-.001	-.016	.002
38	.027	-.002	.003	-.003	.000	.012	.000	-.022	-.004	.005	.005	.000	-.003	.002	.003	-.006	.000	-.037	.006	-.001
39	-.005	.000	.004	-.009	.003	.000	.001	.007	.000	-.003	.002	.004	.004	.003	-.005	-.005	-.001	-.006	-.004	.003
40	.009	.008	.000	-.007	.000	.010	.000	.007	.006	.003	.007	.016	.000	.007	-.001	.001	.001	.000	.000	.002

Tableau 5 (suite)

Matrice des covariances résiduelles pour l'analyse du modèle à deux facteurs

	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39
2																			
3																			
4																			
5																			
6																			
7																			
8																			
9																			
10																			
11																			
12																			
13																			
14																			
15																			
16																			
17																			
18																			
19																			
20																			
21																			
22	-.006																		
23	.008	.016																	
24	-.020	.000	-.013																
25	.009	-.005	.010	.008															
26	.010	-.004	.020	.021	-.023														
27	-.008	.004	-.011	.004	.006	.004													
28	-.002	.011	-.015	-.010	.011	.014	.001												
29	.008	.000	.016	.019	-.011	-.019	.003	.002											
30	.003	-.007	.022	.025	-.005	-.015	.004	.008	-.008										
31	.013	-.007	.010	.010	-.003	-.009	.002	.003	.002	-.059									
32	-.039	.002	.024	.004	.003	-.006	.005	.023	-.011	.002	-.004								
33	.005	-.001	.000	.007	-.007	.002	-.002	-.017	.001	.000	.001	.018							
34	-.037	.007	.021	-.011	.002	-.005	.000	.022	-.004	.000	.002	-.077	.013						
35	.004	-.002	-.028	-.007	.001	.007	-.007	-.007	.012	.012	.005	.019	-.011	.023					
36	-.004	.004	.006	.006	.000	.003	-.013	.000	-.004	-.005	-.003	-.011	.004	-.010	.006				
37	.013	-.004	.008	.008	-.017	-.008	.008	.007	-.007	-.017	-.010	-.004	.004	.002	.012	-.008			
38	-.003	.000	-.029	-.002	.016	.021	-.013	-.014	.012	-.003	-.008	.008	.012	.012	-.018	.009	.020		
39	.008	.004	-.018	-.013	.011	.015	-.003	-.003	.007	.010	.006	.008	-.001	.005	-.010	-.002	.000	-.014	
40	-.010	.002	-.002	-.019	.002	.002	-.004	-.005	.006	.001	.003	-.002	.002	-.008	-.001	.000	.003	-.002	-.014

La deuxième stratégie retenue consiste à examiner la racine carrée de l'erreur quadratique moyenne (root mean square residual). Si cette valeur est égale ou plus petite à celle de l'erreur standard typique des résidus (c.-à-d. $4 \times \frac{1}{\sqrt{n}}$), on peut alors présumer que le modèle proposé est adéquat, c'est-à-dire qu'il n'est pas contaminé par la présence d'autres dimensions. Les résultats de la présente étude indiquent que la valeur de la racine carrée de l'erreur quadratique moyenne (0.0103) est inférieure à la valeur de l'erreur standard typique des résidus (0.1082). Il semble donc que la structure à deux dimensions du BIDR soit confirmée. Les coefficients de saturation associés aux deux facteurs sont rapportés au Tableau 6.

Tableau 6

Coefficients de saturation de l'analyse confirmatoire à deux facteurs du BIDR

Item	Coefficient de saturation du facteur 1	Coefficient de saturation du facteur 2
1	.23	0
2	.38	0
3	.31	0
4	.59	0
5	.56	0
6	.31	0
7	.34	0
8	.51	0
9	.39	0
10	.45	0
11	.36	0
12	.39	0
13	.44	0
14	.65	0
15	.18	0
16	.42	0
17	.42	0
18	.32	0
19	.54	0
20	.49	0
21	0	.46
22	0	.48
23	0	.38
24	0	.42
25	0	.43
26	0	.46
27	0	.29
28	0	.51
29	0	.40
30	0	.29
31	0	.37
32	0	.42
33	0	.51
34	0	.41
35	0	.30
36	0	.32
37	0	.34
38	0	.22
39	0	.41
40	0	.47

L'étape suivante consiste à vérifier si chacune des dimensions postulées par Paulhus respectent le postulat d'unidimensionalité. L'examen de la matrice des covariances résiduelles

pour la dimension autoduperie (Tableau 7) révèle l'absence d'un groupement de valeurs élevées signifiant la présence d'une ou plusieurs autres dimensions. Dans le même sens, les résultats démontrent que la valeur de la racine carré de l'erreur quadratique moyenne (0.0109) est inférieure à celle de l'erreur standard typique des résidus (0.1082). Les coefficients de saturation pour la sous-échelle d'autoduperie sont présentés au Tableau 8.

Tableau 7

Matrice des covariances résiduelles associées au facteur d'autoduperie

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19
2	.05																		
3	-.02	.03																	
4	.04	-.25	.15																
5	.15	-.04	.10	-.19															
6	-.27	.08	-.10	.14	.10														
7	-.02	-.03	.03	.02	-.07	-.05													
8	.19	-.08	-.27	.08	-.13	.03	-.08												
9	.07	-.02	.15	-.22	-.12	.17	-.10	.03											
10	-.14	.00	-.16	.20	.13	-.13	.07	-.15	.12										
11	.02	.06	.04	-.00	-.06	.04	-.18	-.09	-.13	.09									
12	-.01	-.10	-.04	.10	.18	.09	.03	.12	.09	-.08	.05								
13	.03	-.00	.00	.00	-.01	.05	-.05	-.03	-.08	.01	-.08	.06							
14	-.03	.16	.10	.07	-.34	-.26	.08	-.02	.08	-.11	-.02	.03	.02						
15	-.07	.03	-.03	-.03	.04	.01	.09	.10	-.14	.09	-.21	.10	-.07	.04					
16	.04	.08	.02	.03	.15	.12	.06	.10	.05	-.04	.10	-.24	.04	.05	-.11				
17	-.09	.11	-.05	.10	.12	-.08	-.09	.11	.00	.06	.09	-.10	-.02	.06	-.11	-.14			
18	.22	-.09	.07	-.15	-.09	.27	.06	-.21	.08	.07	-.13	-.03	-.08	.12	-.14	-.06	.06		
19	-.09	.09	-.09	.07	.18	-.10	.11	.06	.06	-.13	.05	-.00	.01	.02	.06	-.12	-.14	.11	
20	-.11	-.05	-.01	-.10	.20	.00	.02	.12	-.01	.07	.15	-.19	.07	.11	.07	-.12	-.08	.03	-.18

Tableau 8

Coefficients de saturation associés à la sous-échelle autoduperie

Item	Coefficient de saturation
1	.55
2	.47
3	.27
4	.39
5	.41
6	.43
7	.29
8	.42
9	.38
10	.33
11	.42
12	.54
13	.44
14	.58
15	.23
16	.37
17	.32
18	.17
19	.36
20	.58

Pour la dimension hétéroduperie, des résultats similaires sont observés. Premièrement, la présence d'aucun groupement de covariances résiduelles aux valeurs élevées suggèrent que le modèle à une dimension est adéquat (Tableau 9). Deuxièmement, la racine carré de l'erreur quadratique moyenne (0.0126) est également inférieure à la valeur de l'erreur standard typique des résidus (0.1082). Les coefficients de saturation de la sous-échelle d'hétéroduperie sont présentés au Tableau 10.

Tableau9

Matrice des covariances résiduelles associées au facteur hétéroduperie

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19
2	.03																		
3	.03	.12																	
4	-.17	-.00	-.20																
5	.13	-.06	.04	.07															
6	.14	-.06	.10	.17	.25														
7	-.05	.04	-.15	.03	.06	.03													
8	-.03	.09	-.28	-.16	.06	.06	-.02												
9	.12	-.01	.08	.16	-.03	-.22	.02	-.05											
10	.09	-.06	.19	.26	-.03	-.14	.05	.07	-.07										
11	.19	-.06	.08	.11	-.02	-.08	.03	.02	.03	-.56									
12	-.22	.05	.22	.09	.08	.00	.09	.26	-.05	.09	.03								
13	.06	-.04	-.13	.02	-.11	-.05	-.04	-.28	-.04	-.01	.00	.22							
14	-.20	.11	.21	-.04	.08	.03	.04	.26	.03	.06	.08	-.57	.18						
15	.02	-.04	-.34	-.10	-.01	.02	-.09	-.13	.09	.11	.04	.18	-.16	.24					
16	.01	.05	.04	.07	.02	.04	-.12	.00	-.03	-.03	-.01	-.05	.04	-.03	.05				
17	.16	-.05	.01	.06	-.19	-.11	.07	.00	-.10	-.16	-.09	.00	-.01	.09	.08	-.07			
18	-.04	-.02	-.35	-.05	.13	.16	-.14	-.20	.08	-.03	-.09	.08	.06	.13	-.20	.09	.17		
19	.08	.04	-.23	-.16	.10	.12	-.04	-.08	.05	.10	.06	.10	-.05	.08	-.12	-.02	-.01	-.17	
20	-.04	.03	-.04	-.17	.03	.04	-.02	-.05	.08	.04	.05	.06	.03	.02	-.02	.03	.05	-.02	-.13

Tableau 10

Coefficients de saturation associés à la sous-échelle hétéroduperie

Items	Coefficient de saturation
1	.26
2	.39
3	.28
4	.61
5	.59
6	.33
7	.37
8	.47
9	.39
10	.41
11	.33
12	.42
13	.40
14	.67
15	.16
16	.43
17	.42
18	.28
19	.49
20	.51

Les résultats obtenues de l'analyse de multidimensionalité supportent le modèle à deux dimensions proposé par Paulhus (1984). Ainsi, le BIDR semble composé des facteurs autoduperie et hétéroduperie et chacun d'eux respecte le postulat d'unidimensionalité.

CONCLUSION

Cette recherche portait sur les propriétés métriques du BIDR. À la lumière des résultats obtenus, il semble que ce questionnaire soit composé de deux dimensions, soit l'autoduperie et l'hétéroduperie. L'étude de la structure factorielle à l'aide de la théorie moderne de réponses aux items permet d'évaluer avec plus de précision la dimensionalité de l'échelle. Une des principales forces de ce mémoire réside principalement dans le type d'analyses utilisé pour obtenir plus d'informations sur les propriétés métriques de l'échelle de désirabilité sociale. Les résultats obtenus à partir du DIMTEST et du NOHARM ont permis d'analyser avec plus de précision les réponses des participants.

Le DIMTEST (1987) semble donc être un outil prometteur pour évaluer la dimensionalité d'un test. Ce logiciel est basé sur l'indépendance essentielle et permet d'évaluer les dimensions dominantes. Ce logiciel est basé sur une approche non paramétrique et, partant, élimine les problèmes d'ajustement des données à des modèles paramétriques. Cependant, les approches non paramétriques ne devraient pas, et ne peuvent pas, remplacer des modèles de la TRI plus puissants. Par exemple, l'approche paramétrique semble plus adéquate que l'approche non paramétrique pour l'évaluation des tests de mise en équivalence (test equating), les tests adaptatifs (adaptive tests) de même que les tests visant à évaluer les habiletés cognitives (Embretson, 1994; Kelderman & Rijkes, 1994).

Par ailleurs, l'utilisation de la technique d'analyse factorielle non linéaire (AFNL) pour estimer la dimensionalité d'un test semble prometteuse. Cependant, le logiciel NOHARM est limité puisqu'il ne s'appuie sur aucun test statistique pour juger de la signification des résultats (Berger & Knol, 1990). Des recherches futures devront être menées pour développer une statistique d'ajustement permettant d'évaluer de façon probabiliste les résultats obtenus par le NOHARM.

La reprise de cette étude en utilisant des modèles multidimensionnels de la TRI est suggérée. Toutefois, ce genre d'analyse devrait être précédé d'un examen métrologique du BIDR à partir des plus récentes méthodes d'analyse d'items (p. ex. Ramsay, 1995 ; Thissen, 1986). En effet, quelques items du BIDR semblent ambigus et nuire à la validité du test. De plus, en raison de sa longueur, il est difficile de l'utiliser dans le quotidien.

RÉFÉRENCES

Ackerman, T. (1996). Graphical representation of multidimensional item response theory analyses. *Applied Psychological Measurement*, 20, 311-329.

Ackerman, T., Junker, B., Nandakumar, R., & Stout, W. (1992, mars). *A practitioner's guide for examining test dimensionality*. Papier présenté au Annual Meeting of the American Educational Research Association, Montebello, QC.

Berger, M. (1983). Toward maximizing the utility of consumer satisfaction as an outcome. In M. J. Lambert, E. R. Christensen & S.S. DeJulio (Eds.), *The assessment of psychotherapy outcome* (pp. 56-80). New York: John Wiley & Sons.

Berger, M. P. F., & Knol, D. L. (1990, April). On the assessment of dimensionality in multidimensional item response theory models. Papier présenté au Annual Meeting of the American Educational Research Association, Boston, MA.

Biggs, J. B. (1993). What do inventories of students' learning process really measure? A theoretical review and clarification. *British Journal of Educational Psychology*, 63, 3-19.

Block, J. (1965). *The challenge of response sets*. New York: Appleton-Century-Crofts.

Byrne, D. (1961). The Repression-Sensitization scale: Rationale, reliability, and validity. *Journal of Personality*, 29, 334-349.

Chan, Y-Y., & Watkins, D. (1994). Classroom environment and approaches to learning: An investigation of the actual and preferred perceptions of Hong Kong secondary school students. *Instructional Science*, 22, 233-246.

Christoffersson, A. (1975). Factor analysis of dichotomized variables. *Psychometrika*, 40, 5-32.

Cofer, C. N., Chance, J., & Judson, A. J. (1949). A study of malingering on the MMPI. *Journal of Psychology*, 27, 491-499.

Courmoyer, L., & Sabourin, S. (1991). Autoduperie et Hétéroduperie: facteurs contaminant l'évaluation de la détresse psychologique et de la satisfaction de la clientèle en relation d'aide. *Canadian Journal of Behavioral Science*, 23, 41-52.

Crowne, D. P., & Marlowe, D. (1960). A new scale of social desirability independent of psychopathology. *Journal of Consulting Psychology*, 24, 349-354.

De Ayala, R. J. (1993). An introduction to polytomous item response theory models. *Measurement and Evaluation in Counseling and Development*, 25, 172-189.

De Champlain, A. F., & Gessaroli, M. E. (1991, April). *Assessing test dimensionality using an index based on nonlinear factor analysis*. Papier présenté aux Annual Meeting of the American Educational Research Association, Chicago, IL.

Edwards, A. L. (1957). *The social desirability variable in personality assessment and research*. New York: Dryden Press.

Edwards, A. L., Diers, C. J., & Walker, J. N. (1962). Response sets and factor loading on sixty-one personality scales. *Journal of Applied Psychology*, 46, 220-225.

Embretson, S. E. (1994). Application to cognitive design systems to test development. In C. Reynold (Eds.), *Advances in cognitive assessment; An interdisciplinary perspective* (pp107-135). Plenum

Evans , R. G., & Forbach, G. B. (1982). Intellectual ability correlates of the Marlowe-Crowne Social Desirability Scale. *Journal of Personailty Assesment*, 46, 59-62.

Fraser, C. (1988). NOHARM: A computer program for fitting both unidimensional and multidimensional normal ogive models of latent trait theory [Computer software]. Armidale, Australia: University of New England, Centre for Behavioral Studies.

Fraser, C., & McDonald, R. P. (1988). NOHARM: Least squares items factor analysis. *Multivariate Behavioral Research*, 23, 267-269.

Frenkel-Brunswick, E. (1939). Mechanisms of self-deception. *Journal of Social Psychology*, 10, 409-420.

Gable, R. K., & Wolf, M. G. (1993). *Instrument development in the affective domain: Measuring attitudes and values in corporate and school setting*. (2nd ed.). Boston MA: Kluwer Academic Publishers.

Garfield, S. L. (1983). Some comments on consumer satisfaction in behavior therapy. *Behavior Therapy*, 2, 237-241.

Goldstein, H. (1980). Dimensionality, bias, independence and measurement scale problems in latent trait test score models. *British Journal of Mathematical and Statistical Psychology*, 33, 234-246.

Goldstein, H., & Wood, R. (1989). Five decades of item response modelling. *British Journal of Mathematical and Statistical Psychology*, 42, 139-167.

Gur, R. C., & Sackeim, H. A. (1979). Self-deception: A concept in search of a phenomenon. *Journal of Personality and Social Psychology*, 37, 147-169.

Guttman, L. (1950). The principal components of scale analysis. In S. S. Stouffer (Eds.), *Measurement and prediction* (pp. 312-361). Princeton NJ: University Press.

Guttman, L. (1944). A basis for scaling qualitative data. *American Sociological Review*, 80, 139-150.

Hambleton, R. K., & Rovinelli, R. J. (1986). Assessing the dimensionality of a set of test items. *Applied Psychological Measurement*, 10, 287-302.

Hambleton, R. K., & Swaminathan, H. (1985). *Item response theory: Principles and application*. Boston: Kluwer-Nijhoff.

Hambleton, R. K., Swaminathan, H., Cook, L. L., Eignor, D. R., & Gifford, J. A. (1978). Developments in latent trait theory: Models, technical issues and application. *Review of Educational Research*, 48, 467-510.

Hattie, J. (1985). Methodology review: Assessing unidimensionality of tests and items. *Applied Psychological Measurement*, 9, 139-164.

Hattie, J. (1984). An empirical study of various indices for determining unidimensionality. *Multivariate Behavioral Research*, 19, 49-78.

Hattie, J., Krakowski, K., Rogers, H. J., & Swaminathan, H. (1996). An assessment of Stout's index of essential unidimensionality. *Applied Psychological Measurement*, 20, 1-14.

Hemker, B. T. (1996). *Unidimensional IRT models for polytomous items, with results for Mokken scale analysis*. Unpublished doctoral dissertation, Utrecht University, The Netherlands.

Hemker, B. T., Sijtsma, K., & Molenaar, I. W. (1995). Selection of unidimensional scales from a multidimensional item bank in the polytomous Mokken IRT model. *Applied Psychological Measurement*, 19, 337-352.

Hulin, C. L., Drasgow, F., & Parsons, L. K. (1983). *Item response theory*. Homewood IL: Dow-Jones-Irwin.

Humphreys, L. G. (1986). An analysis and evaluation of test and item bias in the prediction context. *Journal of Applied Psychology*, 71, 327-333.

Humphreys, L. G. (1985). General intelligence: An integration of factor, test, and simplex theory. In B. B. Wolman (Eds.), *Handbook of intelligence* (pp. 201-224). New York: Wiley.

Humphreys, L. G. (1984). *A theoretical and empirical study of the psychometric assessment of psychological test dimensionality and bias* (ONR Research Proposal). Washington, DC: Office of Naval Research.

Humphreys, L. G. (1981). The primary mental ability. In M. P. Friedman, J. P. Das, & N. O'connors (Ed.), *Intelligence and learning* (pp. 87-102). New York: Plenum Press.

Jackson, D. N., & Messick, S. (1962). Response styles on the MMPI: Comparison of clinical and normal samples. *Journal of Abnormal and Social Psychology*, 65, 285-299.

Jacobson, L. I., Kellogg, R. W., Cauce, A. M., & Slavin, R. S. (1977). A multidimensional social desirability inventory. *Bulletin of the Psychometric Society*, 9, 109-110.

Junker, B (1992, April). *Ability estimation in unidimensional models when more than one trait is present*. Papier présenté au Annual Meeting of the American Educational Research Association, San Francisco.

Junker, B. W. (1991). Essential independence and likelihood-based ability estimation for polytomous items. *Psychometrika*, 56, 255-278.

Junker, B. (1990). *Progress in characterizing the monotone unidimensional IRT representation*(Technical Rep. No. 498). Pittsburgh PA: Carnegie Mellon University, Department of Statistics.

Kalliopuska, M. (1992). Social desirability related to social class among adults. *Psychological Reports*, 70, 808-810.

Kelderman, H., & Rijkes, C. P. M. (1994). Loglinear multidimensional IRT models for polytomous scored items. *Psychometrika*, 59, 144-177.

Knol, D. L., & Berger, M. P. F. (1991). Empirical comparison between factor analysis and multidimensional item response models. *Multivariate Behavioral Research*, 26, 457-477

Larsen, K. S., Martin, H. J., Ettinger, R. H., & Nelson, J. (1976). Approval seeking, social cost, and aggression: A scale and some dynamics. *Journal of Social Psychology*, 94, 3-11.

LeVois, M., Nguyen, T. D., & Attkisson, C. (1981). Artifact in client satisfaction assessment. *Evaluation and Program Planning*, 4, 139-150.

Linden, W., Paulhus, D. L., & Dobson, K. S. (1986). Effects of response styles on the report of psychological and somatic distress. *Journal of Consulting and Clinical Psychology*, 54, 309-313.

Lord, F. M. (1980). *Applications of item response theory to practical testing problems*. Hillsdale, NJ: Lawrence Erlbaum.

Lord, F. M., & Novick, M. R. (1968). *Statistical theories of mental test scores*. Reading MA: Addison-Wesley.

Lumsden, J. (1957). A factorial approach to unidimensionality. *Australian Journal of Psychology*, 9, 105-111.

McDonald, R. P. (1985). *Factor analysis and related methods*. Hillsdale NJ: Erlbaum.

McDonald, R. P. (1983). Exploratory and confirmatory factor analysis. In H. Warner & S. Messick (Eds.), *Principles of modern psychological measurement*, Hillsdale, NJ: Lawrence Erlbaum.

McDonald, R. P. (1982). Linear versus nonlinear models in item response theory. *Applied Psychological Measurement*, 6, 379-396.

McDonald, R. P. (1981). The dimensionality of tests and items. *British Journal of Mathematical and Statistical Psychology*, 34, 100-117.

McDonald, R. P. (1979). The structural analysis of multivariate data: A sketch of general theory. *Multivariate Behavioral research*, 14, 21-38.

McDonald, R. P. (1967). Nonlinear factor analysis. *Psychometrika Monograph*, 15.

McDonald, R. P., & Ahlawat, K. S. (1974). Difficulty factors in binary data. *British Journal of Mathematical and Statistical Psychology*, 27, 82-99.

McDonald, R. P., & Mulaik, S. A. (1979). Determinacy of common factors: A nontechnical review. *Psychological Bulletin*, 86, 297-306.

Meehl, P. E., & Hathaway, S. R. (1946). The K factor as a suppressor variable in the Minnesota Multiphasic Personality Inventory. *Journal of Applied Psychology*, 30, 525-564.

Millham, J. (1974). Two components of need for approval score and their relationship to cheating following success and failure. *Journal of Research in Personality*, 8, 378-392.

Millham, J., & Jacobson, L. I. (1978). The need for approval. In H. London & J. E. Exner (Eds.), *Dimensions of personality* (pp. 365-390). New York: Wiley.

Molenaar, I. W. (1997). Nonparametric models for polytomous responses. In W. J. van der Linden & R. K. Hambleton (Eds.), *Handbook of modern item response theory* (pp. 369-380). New York: Springer.

Nandakumar, R. (1994). Assessing dimensionality of a set of item responses--Comparison of different approaches. *Journal of Educational Measurement*, 31, 17-35.

Nandakumar, R. (1993). Assessing essential unidimensionality of real data. *Applied Psychological Measurement*, 17, 29-38.

Nandakumar, R. (1991). Traditional dimensionality vs essential unidimensionality. *Journal of Educational Measurement*, 28, 1-19.

Nandakumar, R. (1987). *Refinements of Stout's procedure for assessing latent trait dimensionality*. Unpublished doctoral dissertation, University of Illinois, Urbana-Champaign.

Nandakumar, R., Stout, W. (1993). Refinements of Stout's procedure for assessing latent trait unidimensionality. *Journal of Educational Statistics*, 18, 41-68.

Paulhus, D. (1991). Balanced inventory of desirable responding. In J.P. Robinson, P.S. Shaver & L.S. Wrightsman. *Mesure of personality and social psychological attitude*. New York: Academic Press.

Paulhus, D. (1987). Attitudes and social cognition: Desirable responding triggered by affect: Automatic egotism? *Journal of Personality and Social Psychology*, 52, 245-259.

Paulhus, D. (1986). Self-deception and impression management in test responses. In A. Angleitner & J. S. Wiggins (Eds.), *Personality assessment via questionnaires: Current issues in theory and measurement* (pp. 143-165). New York: Springer-Verlag.

Paulhus, D. L. (1984). Two-component models of socially desirable responding. *Journal of Personality and Social Psychology*, 46, 598-609.

Paulhus, D. L. (1982). Individual difference, self-presentation, and cognitive dissonance: Their concurrent operation in forced compliance. *Journal of Personality and Social Psychology*, 43, 838-852.

Paulhus, D. L., & Levitt, K. (1987). Attitudes and social cognition: desirable responding triggered by affect: Automatic egotism? *Journal of Personality and Social Psychology*, 52, 245-259.

Porrata, J. L. (1993). Differences in social desirability scores of students in private and public school. *Psychological Reports*, 72, 967-970.

Ramanaiah, N. V., & Martin, H. J. (1980). On the two-dimensional nature of the Marlowe-Crowne social desirability scale. *Journal of Personality Assessment*, 44, 507-514.

Ramsey, J. O. (1995). Testgraf : *A program for the graphical analysis of multiple choice test and questionnaire data*. Unpublished manuscript, McGill University, Montréal.

Reckase, M. D. (1985). The difficulty of test items that measure more than one ability. *Applied Psychological Measurement*, 9, 401-412.

Reckase, M. D. (1979). Unifactor latent trait models applied to multifactor test: Results and implication. *Journal of Educational Statistics*, 4, 207-230.

Sabourin, S., Laferrière, N., Sicuro, F., Coallier, J., Cournoyer, L., & Gendreau, P. (1989). Social desirability, psychological distress and client satisfaction with mental health treatment. *Journal of Counseling Psychology*, 36, 351-356.

Sackeim, H. A., & Gur, R. C. (1978). Self-deception, self-confrontation, and consciousness. In J. Masling (Eds.), *Empirical studies of psychoanalytical theories Vol. 1*, (pp.101-157). Hillsdale, NJ: Erlbaum.

Schaeffer, N. C. (1988). An application of item response theory to the measurement of depression. *Sociological Methodology*, 18, 271-307

Scheirer, M. A. (1978). Program participants' positive perceptions. Psychological conflict of interest in social program evaluation. *Evaluation Quarterly*, 2, 53-69.

Stout, W. (1990). A new item response theory modeling approach with applications to unidimensionality assessment and ability estimation. *Psychometrika*, 55, 293-325.

Stout, W. (1987). A nonparametric approach for assessing latent trait unidimensionality. *Psychometrika*, 52, 589-617.

Stout, W., Habing, B., Douglas, J., Kim, H. R., Roussos, L., & Zhang, J. (1996). Conditional covariance-based nonparametric multidimensionality assessment. *Applied Psychological Measurement*, 20, 331-354.

Stout, W. F., Nandakumar, R., Junker, B., Chang, H. H., & Steindinger, D. (1991). *DIMTEST and TESTSIM* (Computer program). Champaign: University of Illinois, Department of Statistics.

Takane, Y., & De Leeuw, J. (1987). On the relationship between item response theory and factor analysis of discretized variables. *Psychometrika*, 52, 393-408.

Thissen, D. M. (1986). Multilog : Item analysis and scoring with multiple category response model. Moresville, In : Scientific Software.

Traub, R. E. (1983). A priori considerations in choosing an item response model. In R. K. Hambleton (Eds.), *Application of item response theory*. British Colombia: Educational Research Institute of British Colombia.

Traub, R. E., & Lam, Y. R. (1985). Latent structure and item sampling models for testing. *Annual Review of Psychology*, 36, 19-48

Tucker, L. R., & Lewis, C. (1973). A reliability coefficient for maximum likelihood factor analysis. *Psychometrika*, 38, 1-10.

Ullmann, L. P. (1962) An empirically derived MMPI scale wich measure facilitation-inhibition of recognition of threatening stimuli. *Journal of Clinical Psychology*, 18, 127-132.

Walker, D. A. (1931). Answer pattern and score scatter in tests and examinations. *British Journal of Psychology*, 22, 73-86.

Watkins, D. (1996). The influence of social desirability on learning process questionnaires: A neglected possibility? *Contemporary Educational Psychology*, 21, 80-82.

Webb, E. J., Campbell, D. T., Schwartz, R. D., & Sechrest, L. (1970). *Unobtrusive measures: Non-reactive research in the social sciences*. Chicago: Rand McNally.

Wiggins, J. S. (1964). Convergences among stylistic response measures from objective personality tests. *Educational and Psychological Measurement*, 24, 551-562.

Wiggins, J. S. (1959). Interrelationships among MMPI measures of dissimulation under standard and social desirability instructions. *Journal of Consulting Psychology*, 23, 419-427.

Yen, W. M. (1985). Increasing item complexity: A possible cause of scale shrinkage for unidimensional item response theory. *Psychometrika*, 50, 399-410.

Zwick, R. (1987). Assessing the dimensionality of NAEP reading data. *Journal of Educational Measurement*, 24, 293-308.

APPENDICE A

L'approche traditionnelle du modèle de la TRI

Selon les modèles de traits latent, chaque candidat est représenté par un vecteur de valeurs possibles pour les variables θ , où plusieurs candidats peuvent être assignés à chaque θ . Associée avec chaque item i , une fonction de réponses aux items (FRI), $P_i(\theta)$, définit la probabilité qu'un candidat sélectionné aléatoirement parmi un groupe de candidats d'habileté θ réussisse correctement l'item. L'échantillonnage aléatoire de candidats d'une population spécifique est associé à une distribution $F(\theta)$ sur l'espace du trait latent des habiletés (θ) et donc à une distribution pour les réponses du test $\mathbf{U}_N = (U_1, \dots, U_N)$ fournies par des candidats choisis aléatoirement. Le vecteur de réponses aléatoires du test \mathbf{U}_N sert souvent de référence aux tests ou questionnaires. La variable aléatoire U_i référera à l'item i . Quant à \mathbf{u}_N et u_i , ils représentent les valeurs observées de \mathbf{U}_N et de U_i . Pour sa part, $\mathbf{u}_N = (u_1, \dots, u_n)$ constitue une séquence de 0 et de 1. Ainsi, $U_i = 1$ indique une réponse correcte et $U_i = 0$ signifie une réponse incorrecte à l'item i pour n'importe quel candidat choisi aléatoirement. Le vecteur latent aléatoire est dénoté par Θ et des valeurs particulières prises sur Θ seront dénotées par θ . Il faut par conséquent noter que $P_i(\theta) = P[U_i = 1 \mid \Theta = \theta] = E[U_i \mid \Theta = \theta]$ pour tout i, θ . Pour les conventions de la notation, $P(\mathbf{u}_N \mid \theta)$ représente la distribution conditionnelle $P[\mathbf{U}_N = \mathbf{u}_N \mid \Theta = \theta]$. Il est important d'insister sur le fait qu'un test \mathbf{U}_N peut être associé à plusieurs modèles possibles de trait latent. Il y a donc plusieurs choix possibles de paires $F(\theta), P(\mathbf{u}_N \mid \theta)$ de façon telle que pour tout \mathbf{u}_N

$$P[\mathbf{U}_N = \mathbf{u}_N] = \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} P(\mathbf{u}_N \mid \theta) dF(\theta)$$

Il est à noter que (\mathbf{U}_N, Θ) sont des paires de vecteurs aléatoires pour lesquelles la distribution jointe est spécifiée par la distribution marginale $F(\theta)$ et la distribution conditionnelle $P(\mathbf{u}_N | \theta)$. Les modèles latents pour \mathbf{U}_N seront dénotés par $(\mathbf{U}_N, \Theta, P(\mathbf{u}_N | \theta), F(\theta))$ ou tout simplement par (\mathbf{U}_N, Θ) . Trois caractéristiques des modèles latents doivent être signalées:

1. Le modèle (\mathbf{U}_N, Θ) est dit monotone (M) si $P[U_i = 1, \dots, U_{i_k} = 1 | \Theta = \theta]$ est non-décroissante en θ pour chaque (U_i, \dots, U_{i_k}) de \mathbf{U}_N (ici, par définition, $\theta_1 = \theta_2$ si et seulement si $\theta_{1i} = \theta_{2i}$ pour chaque coordonné i ; ainsi que $1 = i_1 < \dots < i_k = N$). Le modèle est donc strictement monotone si le terme non-décroissant est remplacée par le terme strictement croissant.

2. Le modèle (\mathbf{U}_N, Θ) possède d dimension si Θ est un vecteur aléatoire de dimensions d . L'habileté de dimensions d est donc dénotée par $(\theta_1, \dots, \theta_d)$. La dimensionalité de Θ sera dénotée par $\dim(\Theta)$ ou d .

3. Le modèle (\mathbf{U}_N, Θ) est dit être localement indépendant (LI) si

$$P(\mathbf{u}_N | \theta) \equiv P[U_1 = u_1, \dots, U_N = u_N | \Theta = \theta] = \prod_{i=1}^N P[U_i = u_i | \Theta = \theta] = \prod_{i=1}^N P_i(\theta)^{u_i} [1 - P_i(\theta)]^{1-u_i}$$

pour tout θ et chacun des 2^N choix de (u_1, \dots, u_N) .

Les modèles de classes latentes les plus communément utilisés sont les modèle LI, M, $d=1$. Dans ce cas M est équivalent aux FRI qui sont toutes monotones. Habituellement pour ces

modèles lorsque M et $d=1$ sont respectés, les FRI sont strictement monotones. À noter que dans le cas LI, $d=1$ possédant une distribution latente, $F(\theta)$ possède une densité $f(\theta)$. En combinant l'équation (1) et (2) pour produire l'équation du modèle usuel de la TRI

$$P[\mathbf{U}_N = \mathbf{u}_N] = \int_{-\infty}^{\infty} \left\{ \prod_{i=1}^N P_i(\theta)^{u_i} [1 - P_i(\theta)]^{1-u_i} \right\} f(\theta) d\theta$$

APPENDICE B

DIMTEST

L'indice statistique utilisé pour estimer l'unidimensionnalité est basé sur le principe d'indépendance essentielle. Selon l'hypothèse nulle qu'un test est unidimensionnel, n'importe quel sous-ensemble de répondants avec des scores approximativement égaux sur un test composé d'un nombre très grand d'items devraient approximativement posséder la (les) mêmes habiletés. Donc ceci implique que l'hypothèse d'indépendance locale est éprouvée approximativement.

Les 4 étapes de DIMTEST

Étape 1. Un ensemble de M items est choisis parmi les items du test pour qu'il soit le plus unidimensionnel possible. Cet ensemble sera appelé le Sous-ensemble d'Estimation 1 (AT1). Il y a trois stratégies pouvant servir à l'identification de ces M items: (1) une analyse de jugement ou d'experts, (2) une solution en composante principale des matrices de corrélations tétrachoriques sur tous les N items, (3) ou une analyse par grappe (Roussos, Stout, & Marden, 1993).

Étape 2. Un deuxième ensemble de M items est choisi parmi les items restants de façon à ce qu'ils soient similaires en regard de leur niveau de difficulté et qu'ils mesurent la même dimension que les items de l'ensemble AT1. Ce groupement sera appelé le Sous-ensemble d'Estimation 2 (AT2). Ce nouveau sous-ensemble sera plus tard utilisé pour corriger les biais de la statistique T. Les biais proviennent du fait que la moyenne augmente pour de petits tests. Étant une conséquence de la sélection d'items qui sont trop homogènes en respect du niveau de difficulté.

Étape 3. Les items restants $n=N-2M$ sont compris dans l'ensemble où les candidats sont évalués et par la suite répartie en sous-groupes selon leur score total à ces items, appelé Sous-ensemble de Partitionnement (PT). Parce que le nombre de candidats dans chaque groupe de sous-ensemble de score est typiquement trop petit, Stout (1987) a recommandé que les répondants soient assignés aux groupes sur la base du score total pour qu'un large nombre (approximativement 20) soient dans chaque groupe.

Étape 4. Pour chacun de ses sous-groupes, la variance estimée ($\hat{\sigma}_k^2$, $\hat{\sigma}_{U,k}^2$) et l'estimé de l'erreur standard (S_k) sont calculés en utilisant les items de AT1 (voir Nandakumar & Stout, 1993, Équation 3). L'estimé de la variance usuelle $\hat{\sigma}_k^2$ est la variance observée du sous-ensemble AT1. De plus, il est sensible au non-respect du postulat d'unidimensionalité. L'estimé de la variance unidimensionnelle $\hat{\sigma}_{U,k}^2$ correspond à la somme des variances pour les k groupes et reste le même en regard du nombre de dimensions du test. Ces estimés sont donc additionnés pour les k sous-groupes :

$$T_L = \frac{1}{\sqrt{K}} \sum_{k=1}^K \frac{|\hat{\sigma}_k^2 - \hat{\sigma}_{U,k}^2|}{S_k}$$

Une statistique similaire, T_B est calculée pour les items dans AT2. T_L est une mesure de multidimensionalité dans chaque sous-groupe k . Si le postulat d'unidimensionalité est respecté pour le sous-groupe de k candidats alors T_L devrait être égal à zéro, excepté pour l'erreur statistique. Ceci est identique à subdiviser les candidats en sous-ensemble selon leur score et

donc, à vérifier si le principe d'indépendance local est respecté. T_L est calculé pour les M items dans AT1 et est sensible à la dimensionalité et aux sources de biais; T_B est basé sur un ensemble similaire dans AT2 et est sensible aux sources de biais mais non à la dimensionalité. Donc, T_B est utilisé pour corriger le biais. La statistique finale est

$$T = \frac{T_L - T_B}{\sqrt{2}}$$

Le principe de base du DIMTEST est que si le principe d'unidimensionalité est respecté alors l'hypothèse d'indépendance essentielle tient approximativement entre chaque sous-groupe de candidats. Aussi, les deux estimés de variance entre les sous-groupes doivent être approximativement égaux. Stout (1987) a démontré que la statistique \underline{T} possède une distribution asymptotiquement normale lorsque le postulat d'unidimensionalité est respecté.

APPENDICE C

NOHARM

NOHARM (Fraser, & McDonald, 1988) est un algorithme destiné à estimer l'ajustement de données à des modèles normaux ogive unidimensionnels et multidimensionnelle de la théorie des traits latents (en utilisant la meilleure approximation en terme de la série polynomiale de Hermite-Tchebycheff). Cette technique utilise les analyses harmoniques (Fraser, 1981; McDonald, 1982).

Considérons premièrement le modèle ogive unidimensionnelle, en utilisant la notation de Lord (1980), le modèle est exprimé comme:

$$P\{y_j = 1|\theta\} = c_j + (1 - c_j)N\left[a_j(\theta - b_j)\right] \quad (1)$$

où θ = est la variable du trait latent
 a_j = est le paramètre de discrimination
 b_j = est le paramètre de difficulté
 c_j = est le paramètre de pseudo-chance
 $N[.]$ = est la fonction de distribution normale

Si nous supposons $f_{j0} = -a_j b_j$ et, $f_{j1} = a_j$ nous pouvons réécrire l'équation (1) comme:

$$P\{y_j = 1|\theta\} = c_j + (1 - c_j)N\left[f_{j0} + f_{j1}\theta\right] \quad (2)$$

L'équation (2) peut alors être généralisée au modèle multidimensionnel :

$$P\left\{y_j = 1 \mid \underline{\theta}\right\} = c_j + (1 - c_j) N\left[\underline{f}_{j0} + \underline{f}_j' \underline{\theta}\right] \quad (3)$$

où $\underline{\theta}$ est un vecteur de trait latent $m \times 1$ et

\underline{f}_j est un vecteur de coefficient $m \times 1$

Soit F une matrice exprimée tel que:

$$F = \begin{bmatrix} \underline{f}'_1 \\ \underline{f}'_2 \\ \vdots \\ \underline{f}'_n \\ \vdots \end{bmatrix} = \begin{bmatrix} f_{11} & f_{12} & \cdots & f_{1m} \\ f_{21} & f_{22} & \cdots & f_{2m} \\ \cdots & \cdots & \cdots & \cdots \\ f_{n1} & f_{n2} & \cdots & f_{nm} \end{bmatrix} \quad (4)$$

et soit P , la matrice de covariances (ou de corrélation) de $\underline{\theta}$.

Le programme calcule les covariances résiduelles des items, et ce après l'ajustement du modèle. De plus, il fournit la racine carré de la moyenne des covariances comme une mesure globale du mauvais ajustement du modèle aux données, de même que le nombre de dimensions et les coefficients de la matrice du modèle postulé. La matrice résiduelle donne une indication du niveau de satisfaction du principe d'indépendance locale du modèle postulé jusqu'au moment de deuxième ordre (McDonald, 1981). Si la racine carré de l'erreur quadratique moyenne est de l'ordre de l'erreur standard typique des résidus (quatre fois l'inverse de la racine carré de la grandeur de l'échantillon), ceci fournit une indication que le nombre postulé de dimensions d'un

test est adéquat. La matrice des covariances résiduelles qui présente des résidus élevés permet d'identifier un groupement d'items qui ne s'ajustent pas au modèle.

Paramétrisation du modèle de facteur commun

Le modèle de trait latent peut être obtenu de celui de Christoffersson (1975) sous la forme suivante:

$$P\left\{y_j = 1 \mid \theta\right\} = c_j + (1 - c_j) N\left[\left(t_j + h_j' \theta\right) / m_j\right] \quad (5)$$

$$\text{où } t_j = f_j 0 / s_j \quad (6)$$

$$h_j = f_j / s_j \quad (7)$$

$$m_j^2 = 1 / s_j^2 \quad (8)$$

$$\text{et } s_j = \text{la racine carrée de } \left(1 + f_j' P f_j\right) \quad (9)$$

en supposant que pour chaque réponse à un item, il y a une variable théorique v , laquelle s'ajuste au modèle de facteur commun

$$v_j = h_j' \theta + e_j$$

avec $\text{Var}(e_j) = m_j^2$ (variance unique).

Nous supposons, de plus que la réponse à l'item y_j égale 1 (correct), si et seulement si v_j est plus grand ou égal à t_j , qui correspond au paramètre de seuil (threshold value).

Procédure d'estimation

La procédure pour l'estimation des paramètres des matrices F et P, et pour le vecteur f_j est une variante de celle décrite dans McDonald (1982). En réécrivant l'équation (3), le modèle est :

$$\hat{y}_j = c_j + (1 - c_j) N \left[f_{j0} + f_j' \theta \right] \quad (10)$$

laquelle sera approximée par une série polynomiale à quatre termes :

$$\hat{y}_j = b_{j0} h_0(z_j) + b_{j1} h_1(z_j) + b_{j2} h_2(z_j) + b_{j3} h_3(z_j) \quad (11)$$

où $z_j = f_j' \theta / d_j$

$d_j = \text{la racine carrée de } \left(f_j' P f_j \right)$

$h_k()$ est la série polynomiale normalisée de Hermite-Tchebycheff à k degrés, obtenue à partir de :

$$h_0(x) = 1$$

$$h_2(x) = (x^2 - 1)/\sqrt{2}$$

$$h_1(x) = x$$

$$h_3(x) = (x^2 - 3x)/\sqrt{8}$$

Les coefficients b_{jk} sont obtenus quant à eux à partir des transformations suivantes :

$$b_{j0} = c_j + (1 - c_j)N[f_{j0} / e_j]$$

$$b_{j1} = (1 - c_j) n(f_{j0} / e_j) (d_j / e_j)$$

$$b_{j2} = (1 - c_j) (f_{j0} / e_j) n(f_{j0} / e_j) (d_j / e_j)^2 / \sqrt{2}$$

$$b_{j3} = (1 - c_j) ((f_{j0} / e_j)^2 - 1) n(f_{j0} / e_j) (d_j / e_j)^3 / \sqrt{6}$$

où d_j est défini plus haut, $e_j =$ est la racine carré de $(1+d_j)^2$ et $n()$ est une fonction de densité normale.

Si nous utilisons l'approximation de l'équation (11), il s'ensuit que la proportion de candidats réussissant l'item j est approximé par b_{j0} , et la proportion que les candidats obtiennent un succès aux items j et k est estimée par

$$\hat{a}_{jk} = b_{j0} b_{k0} + b_{j1} b_{k1} \left(\frac{f_j' P f_k}{d_j d_k} \right) + b_{j2} b_{k2} \left(\frac{f_j' P f_k}{d_j d_k} \right)^2 + b_{j3} b_{k3} \left(\frac{f_j' P f_k}{d_j d_k} \right)^3$$

Le but est de minimiser le carré des différences (squared differences) entre les proportions de l'échantillon a_{jk} , de candidats réussissant les items j et k , et \hat{a}_{jk} défini ci-haut. La minimisation est donc effectuée en respect des paramètres présents dans les matrices F et P . Dans l'ordre pour effectuer cette démarche, on suppose que a_{jj} sert d'approximation pour b_{j0} . Lorsqu'une fonction minimum est trouvée, f_{j0} est donc recalculée.

Cette méthode est considérée comme une procédure raisonnablement robuste ayant pour but l'ajustement du modèle ogive normale multidimensionnelle pour de grands échantillons de données. Dans sa forme générale, cette conception se rapproche de celle présentée par

Christoffersson (1975). La principale différence provient de l'utilisation des moindres carrés, Christoffersson, quant à lui, emploie la méthode des moindres carrés généralisées. Elle se justifie aussi par l'utilisation d'une série de quatre termes comparativement à une série de dix termes pour Christoffersson. La justification d'utiliser une série polynomiale plus courte provenant à l'origine de l'analyse harmonique pour le modèle ogive normal (McDonald, 1967). Christoffersson (1975) a obtenu des résultats mathématique similaires à la solution du problème d'intégration de la série polynomiale. La théorie suppose que le trait latent possède une distribution normale, et de plus des études de simulation avec la technique Monte Carlo, suggèrent que cette méthode soit aussi robuste contre la violations d'une telle hypothèse.

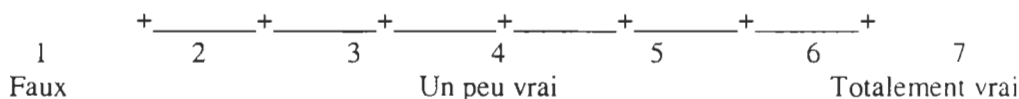
Procédure de minimisation

La fonction à minimiser est la fonction des moindres carrés. Cette fonction effectue la minimisation en utilisant un algorithme de minimisation du modèle quasi-Newton ou du modèle du gradient conjugué (Le choix est fait par le logiciel, tout dépendant du format du problème et de la quantité de mémoire disponible).

ANNEXE D

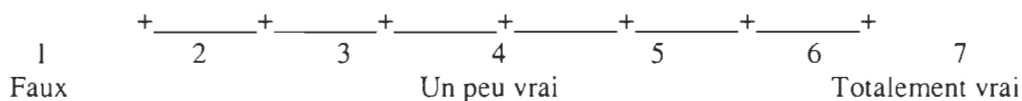
BIDR Version 6 --- Forme 40

Servez-vous de cette échelle de valeurs et inscrivez un chiffre à côté de chaque énoncé pour indiquer à quel point vous êtes d'accord.



- ___1. En général, la première impression que me laissent les gens s'avère juste.
- ___2. Il me serait difficile de me défaire de n'importe laquelle de mes mauvaises habitudes.
- ___3. Il m'importe peu de savoir ce que les gens pensent vraiment de moi.
- ___4. Je n'ai pas toujours été honnête envers moi-même.
- ___5. Je sais toujours pourquoi j'aime quelque chose.
- ___6. Lorsque mes émotions sont sollicitées, mon jugement est affecté.
- ___7. Une fois que je me suis décidé(e), on peut rarement me faire changer d'opinion.
- ___8. Au volant, je deviens dangereux lorsque j'excède la limite de vitesse.
- ___9. Je suis maître(sse) de mon destin.
- ___10. Il m'est difficile de faire abstraction d'une pensée qui me trouble.
- ___11. Je ne regrette jamais mes décisions.
- ___12. Je perds parfois de bonnes occasions parce que je prends trop de temps à me décider.
- ___13. Je vote parce que mon vote peut faire la différence.
- ___14. Mes parents n'étaient pas toujours justes lorsqu'ils me punissaient.
- ___15. Je suis une personne complètement rationnelle.
- ___16. J'accepte rarement les critiques.
- ___17. J'ai énormément confiance en mon jugement.
- ___18. J'ai parfois douté de mes capacités en tant qu'amant(e).
- ___19. Ça me laisse indifférent(e) que certaines personnes ne m'aiment pas.
- ___20. Je ne comprends pas toujours les raisons qui me poussent à faire les choses que je fais.

Servez-vous de cette échelle de valeurs et inscrivez un chiffre à côté de chaque énoncé pour indiquer à quel point vous êtes d'accord.



- ___21. Parfois je mens, s'il le faut.
- ___22. Je ne cherche jamais à dissimuler les erreurs que j'ai commises.
- ___23. Il m'est arrivé(e) de profiter de quelqu'un.
- ___24. Je ne jure jamais.
- ___25. J'essaie parfois de me venger plutôt que de pardonner et d'oublier.
- ___26. J'obéis toujours aux lois, même s'il est peu probable que je me fasse prendre.
- ___27. J'ai parlé en mal d'un(e) ami(e) dans son dos.
- ___28. Lorsque je surprends une conversation privée, j'évite d'écouter.
- ___29. Un(e) caissier(ère) m'a remis trop de monnaie et je ne le lui ai pas mentionné.
- ___30. Je déclare toujours tout aux douanes.
- ___31. Il m'arrivait parfois de voler quand j'étais jeune.
- ___32. Je n'ai jamais jeté de déchets dans la rue.
- ___33. Lorsque je conduis, je dépasse parfois la limite de vitesse.
- ___34. Je ne lis jamais des livres ou des revues érotiques.
- ___35. J'ai fait des choses dont je ne parle pas aux autres.
- ___36. Je n'utilise jamais des choses qui ne m'appartiennent pas.
- ___37. J'ai pris des congés de maladie au travail ou à l'école, même si je n'étais pas vraiment malade.
- ___38. Je n'ai jamais endommagé un livre de bibliothèque ou des articles de magasin sans le signaler à un responsable.
- ___39. J'ai quelques très mauvaises habitudes.
- ___40. Je ne fais pas de commérage au sujet des affaires des autres.

Questionnaire conçu par D.L. Paulhus, 1989. Traduit et adapté par S. Sabourin et Y. Lussier.